Ph.D. THESIS
Essays in the Economic Theory of Organizations

Daniel García González

Universidad Carlos III de Madrid

Departamento de Economía

Supervised by Antonio Cabrales Goitia.

June, 2012

Contents

Introduction

Resumen

Mi tesis doctoral consta de tres capítulos: *Reputation with Endogenous Monitoring*, *Strategic Communication in Networked Organizations* y *Information Technology, Decision Making and Incentives*. En ellos utilizo diferentes herramientas de la Teoría Microeconómica para entender cómo las empresas, organizaciones u otras instituciones ajenas al mercado agregan la información obtenida por sus miembros en su toma de decisiones.

En el primer capítulo, *Reputation with Endogenous Monitoring*, estudio la interacción estratégica entre un individuo que debe tomar decisiones y un experto que posee la información relevante. El experto puede tener preferencias distintas que las del decisor (y esto es su información privada) y sus recomendaciones son estratégicas. Por su parte, el decisor puede verificar la recomendación del agente pero esto es costoso para él. Asumiendo que ambas partes interactúan repetidamente, analizo los incentivos que tiene el decisor para verificar la recomendación del agente dependiendo de su horizonte temporal. Muestro que si el decisor es paciente y está suficientemente seguro que el agente tiene distintas preferencias que él, verificará con menor probabilidad en cualquier equilibrio de Markov que en el caso en que es absolutamente impaciente. Además como al decisor paciente le gustaría comprometerse a verificar con mayor probabilidad de la que resulta del equilibrio, está dispuesto (en algunos casos) a delegar la decisión en uno impaciente. Finalmente, relaciono estos hallazgos teóricos con evidencia empírica obtenida en diversas relaciones bancarias.

En el segundo capítulo, *Strategic Communication in Networked Organizations* analizo organizaciones compuestas por un número arbitrario (finito) de individuos que deben tomar una decisión. Su pago depende de su nivel de coordinación y de la realización de una variable aleatoria sobre la que solamente tienen una señal imperfecta. Mi contribución es que les permito enviar mensajes a otros individuos antes de tomar la decisión. En concreto, cada individuo puede comunicarse con un número determinado de ”vecinos” con los que está unido en la red. Muestro que la necesidad de coordinación y el uso de información imperfecta genera problemas en la comunicación excepto en el caso que las redes generen una distribución de la información igualitaria (por ejemplo, redes ”regulares” en las que todos los individuos tienen el mismo número de vecinos). En otro caso, la comunicación añadirá ruido a la información y la organización no podrá coordinarse eficientemente. Muestro que estas redes igualitarias pueden ser ”redes de equilibrio” si permitimos a los individuos decidir cuánta información adquiren y que los resultados cualitativos pueden extenderse al caso en que la comunicación se produce en rondas sucesivas. La conclusión general del artículo es que en organizaciones en la que la toma de decisiones sea descentralizada, las jerarquías y otras redes cuya distribución de información con desigualdad producen sesgos y ruido en la comunicación e ineficiencia en la toma de decisiones. Este resultado es notablemente distinto del obtenido por modelos de organizaciones en las que la comunicación es no estratégica.

Por último, en el tercer capítulo, *Information Technology, Decision Making and Incentives* analizo el funcionamiento de una organización compuesta por múltiples unidades locales (o individuos) relacionados a través de una unidad central. La unidad central toma una decisión común para todas, dependiendo de la información recogida y transmitida por cada unidad. La transmisión de información se realiza mediante la una señal costosa en forma de esfuerzo. En este contexto, determino la estructura óptima de información que la unidad central obtiene sobre el comportamiento de las unidades locales para tomar decisiones acertadas y encauzar correctamente sus esfuerzos. Encuentro que la cantidad de información que la unidad central quiere adquirir dependerá del grado de volatilidad del entorno y de la importancia de coordinación, pero que nunca deseará observar perfectamente las acciones realizadas por las unidades locales. Finalmente, relaciono estos resultados teóricos con la forma organizativa y la inversión en tecnologías de la información realizada en diversas empresas.

Summary

My Ph.D. thesis contains three Chapters: *Reputation with Endogenous Monitoring*, *Strategic Communication in Networked Organizations* and *Information Technology, Decision Making and Incentives*. In each of them I use different theoretical techniques to understand how firms, organizations and other institutions aggregate the information acquired by their member in order to make appropriate decisions.

In the first Chapter, *Reputation with Endogenous Monitoring*, I analyze the strategic interaction between an uninformed decision-maker and an informed expert. The expert may have a conflict of interest with the decision-maker, but this is his private information. The decision-maker is allowed to verify the recommendation of the expert, but this is costly. Assuming that these two individuals interact repeatedly over time, I study the incentives of the decision-maker to verify the recommendation depending on her time horizon. I show that if the decision-maker is patient and sufficiently pessimistic about the preferences of the agent, she will verify with less probability in any Markov Perfect Equilibrium as compared with the case in which she is myopic. I also show that, due to a lack of commitment, the principal could increase his payoff by delegating to a myopic one. Finally, I relate this theoretical findings with the empirical evidence from the banking industry.

In the second Chapter, *Strategic Communication in Networked Organizations* I analyze organizations comprised of a finite number of individuals, each of whom must take a single decision. Their payoff depends both on their coordination and on the realization of a random variable over which they only observe a noisy signal. I allow them to send messages to other individuals through an undirected network. I show that the need for coordination and the use of partial information difficulties communication even if there is no intrinsic conflict in preferences. I characterize the set of networks that guarantee efficient information transmission, and show that they imply that the ex-post distribution of information is very equal. Otherwise, the communication process will add noise and hamper the quality of the decisions made and the degree of coordination across individuals. I also show that, under some conditions, these networks can be rationalized as the outcome of a (larger) game in which the individuals do also decide how much information to acquire. The main insight of the paper is that in organizations relying on decentralized decision-making hierarchies and other networks generating unequal information distributions create biases and noises in communication, and therefore, inefficiencies in decision making. This result is in sharp contrast with the literature on non-strategic communication in organizations.

Finally, in the third Chapter, *Information Technology, Decision Making and Incentives* I study the behavior of a multi-unit organization where the Headquarters must take a common decision, based on the information transmitted by each unit. The information transmission process takes the form of a signalling game. I analyze the optimal information structure that the Headquarters would choose in order to balance the trade-off between improved decisions from better information and worse incentives. I show that this information structure will depend on the volatility of the environment and the need for coordination, but it will never be the case that the Headquarters want to know perfectly the amount of effort that each local unit makes. I relate these findings to the organizational form and the empirical

evidence on the value of the investments in Information and Communication Technologies.

Dedicatoria

Para mis padres, por estar detrás de todo lo bueno y quererme a pesar de todo lo malo que he hecho.

Para Jesús.

Para Cova, porque sin ella nunca hubiera empezado y, sobre todo, jamás hubiera terminado.

Acknowledgements

> 'It was the best of times, it was the worst of times'
> Charles Dickens

It is time to recap! Five and a half years! Quicker said than done. Five years for seventy-something pages full of extremely boring theoretical results nobody cares about? Is that all[1]? yeah[2]!

Anyhow, this is the end and it's time to say 'Thank you!' I have been extremely lucky to have Antonio as my supervisor. I've learned a lot from him, from the way he thinks and the way he asks. He's extremely supportive in the bad times and got me in the right track in the best times. Life looked always brighter at his office.

I have been not less lucky to have met Marco. I'll never be able to pay him back for so many hours. For having an opinion about every single topic. For sharing it with all of us. For having more faith in me than myself. For teaching me more economics than I can learn. For improving my teaching skills, my grammar skills, my dressing skills...For teaching me almost everything that can be taught [3].

I must thank also to other professors who have been particularly important throughout these years. Luis Garicano for helping me when I was starting. Belén and David, for teaching me the value of math and for all their support. Nezih and Matthias, for caring so much. Luis Corchón, Angel, Misha for all their comments.

Thank you Luis. For being such a good friend. For being always nice when I was not. For smiling even when things looked awful. For so many *cafés* and so many *pulguitas*. For all the laughter and all the pain we went through together. I'm glad things are getting better for you! Thank you Juan. For being such a good friend. For caring so much about me, and being an older brother to me when I needed one. For all those *cañas* and all those *tapas*. For all those *asados* (even those you did when I was away...) Thank you! Thanks to Joaco y Dolo. For being there. For caring so much. For helping me by letting me help you (a bit). Thanks to Lu, for being more crazy than I am and for being such a good friend. Thanks to Eva, for listening. Thank you Zoe, for talking. Thank you Jesper, for reading all these papers. For

[1] Note to self: Do not write Acknowledgements for a Thesis in the midst of an end-of-youth depression

[2] Note to others: Keep reading, it's not that bad

[3] but recall 'nothing that is worth knowing can be taught'

being such a nice guy in Malaga and Chicago. For being so stubborn about our conversation topics. Thank you Efi[4], for cheering us up. Thank you Lián and Luis. Thanks to Edu, Jose and (specially) Carlos for enduring me as flatmate. That must have been a pain.

Finally a customary disclaimer. I warn the reader that those mentioned above are only responsible for the (few) good things that follow, while I am the sole responsible for the (many) errors that remain.

[4]this is not the AER, but hey, you made it!

CHAPTER 1

Reputation with Endogenous Monitoring

'It is with love as with cuckoldom'—the suffering party is at least the third, but generally the last who knows anything about the matter.' Sterne, *Tristram Shandy*

1.1. Introduction

Many economic organizations rely on the long-run concerns of their members to sustain cooperation and reduce the impact of opportunism. Firms may outperform anonymous markets whenever contract enforcement is difficult or costly, since they can identify those who take inefficient actions[1]. For instance, in environments with adverse selection, patient players may be able to build a reputation by mimicking the behavior of cooperative types or by signaling away from defectors [2]. They trade-off short-run gains against future losses that occur whenever their partners identify their behavior as *bad*, and get to expect this bad behavior to be maintained in the future.

Therefore, for reputations to emerge, playing partners must be able to identify *bad* behavior, and the performance of the organization is effectively determined by its monitoring technology. Accordingly, the literature has devoted most of its attention to the identification of conditions in the information structure that result in efficient outcomes. This information structure is, however, taken to be exogenous and independent of the behavior of the individuals. Casual empiricism, however, suggests that in many cases performance measures are not readily observable and costly resources must be devoted to identify potential misbehavior. Therefore, incentives must be in place for players to monitor intensively their partners, identify and deter defectors. In this paper, I am interested in such situations and analyze the impact of this assumption on the outcomes of the interaction[3].

[1] An excellent review is presented in Mailath and Samuelson (2006)

[2] Seminal contributions are Kreps and Wilson (1982) and Milgrom and Roberts (1982), where a firm may be able to credibly threat to predate entrants. Benabou and Laroque (1992) study a model in which a financial advisor is able to convey information credibly. Bar-Isaac (2003) studies a market equilibrium with firms providing high quality, even if this is unverifiable.

[3] To the best of my knowledge there are only two papers with endogenous monitoring in infinite-horizon games. Liu (2011) studies a dynamic game with short-run

More specifically, I study a very simple, infinite-horizon game between two patient players. In each period, one of the players (the principal) has to choose between two alternatives. The principal is ex-ante indifferent between them. However, his payoff of choosing each alternative depends upon the realization of the state of nature, which is drawn anew in every period. The principal is uninformed about this realization but has access to the report of an agent who has some information regarding this realization. The agent's interests, however, may be misaligned with those of the principal and this is his private information. A *good* agent is "committed" to telling the truth every period, while a *bad* agent is strategic and biased towards one alternative. The principal may verify the report of the agent by paying some positive cost.

I concentrate the analysis on stationary Markov Perfect Equilibria (MPE) where each player's strategy is a measurable function of payoff-relevant components of the history. I show that whenever the principal is sufficiently pessimistic about the type of the agent, in any MPE of the infinite-horizon game she monitors less often and the agent lies more often if the principal is a long-run player (compared with a situation where the principal is myopic). The intuition is simple. The strategic type of the agent has incentives to build a reputation for being truthful, in order to get his preferred decision more often in the future. This reputation is constructed by mimicking a truthful, non-strategic type. Reputation is valuable for the principal since it improves the informativeness of the report and saves on monitoring costs. However monitoring is also costly because it may bring bad news, i.e., the agent may be revealed to be strategic and continuation equilibrium yields lower value to the principal. This implies that for a patient principal monitoring intensity becomes lower and, consequently, the strategic type lies more often in equilibrium.

I argue that these effects may be present in a wide variety of economic interactions, where a combination of adverse selection and moral hazard is present. In those environments, monitoring is a fundamental economic activity and the incentives to undertake it crucially depend on the time-horizon of the relationship. According to the conventional view, in a long-term relationship, agents will acquire more information since they can use it in the future[4]. However, monitoring not only yields information but changes the nature of the relationship and constraints the behavior of the monitored agent. This, in turn, may affect the incentives to monitor and the value of the relationship. Hence, by reducing the effort in monitoring activities, a

buyers who may acquire information about previous trades of a seller. Kandori and Obara (2005) who prove a sort of folk theorem in a repeated game where every player may devote effort to monitor others.

[4]See Fama (1985), Diamond (1991)

patient principal can reduce the risk of discovering a bad type and losing a valuable relationship[5].

For instance, relationship finance is commonly assumed to reduce conflicts of interest and increase the effort that banks put in monitoring the actions of their borrowers. This is because banks appropriate the returns of their information in the future and therefore invest more today in acquiring it. This assumes, however, that those returns are positive, which may not be the case in the game I present. Therefore, whenever the mechanism presented in this paper is at work, banks entering in arm's length financing may provide more intensive monitoring than those engaging in relationship finance.

Similar examples abound. Parents may not have the appropriate incentives to monitor whether their kids do their homework, start smoking or get into troubles. Married individuals may face similar problems when monitoring their potentially unfaithful partner. Finally, supranational authorities may struggle to monitor intensively countries under their authority. Notice that in all these situations, the principal has a long-run concern and very limited ability to choose or replace the agent.

More generally, I argue that this paper provides a rationale for why *fraudulent types* survive longer than external observers would expect in many organizations. As an example consider the case of double spies. Intelligence Agencies are unsure whether their spies are double agents and would like to crosscheck their reports. Nevertheless, if every report is required to be checked, the agent is of no value. Even more, if an agent is known to be a double agent, he is useless and the agency may not have access to similar agents or sources of information. Therefore, double agents would survive longer than expected. This was the case of Juan Pujol, a Spanish spy working for both UK and Germany. Although, from the point of view of historical evidence, he had given enough evidence to the Germans that he was a double agent, they maintained their confidence in his reports until the very end[6].

In each of these situations, long-term incentives for the principal may fail to yield higher incentives to monitor. Therefore, one would expect that in organizations where this problem is severe, short-run incentives are used to encourage information acquisition. Indeed this is the case in the case analyzed in a recent contribution (Hertzberg et al. (2010)). They document agency problems related with the time-horizon of relationships between the loan-officer of a bank and the borrowing firm. They show that, whenever the relationship of a

[5]See Subsection 1.4 for a discussion on the assumptions required for the result.

[6]Indeed, Juan Pujol was key in the cover up of the D-Day. More information about him is to be found in Andrew (2009)

given officer with a firm is about to end, the officer sends more accurate reports, and these reports are more likely to include bad news. In order to alleviate these agency problems, the bank has introduced policies of rapid turnover and task reallocation, and compensation packages depending on short-term objectives.

This model may also be interpreted as an expert-agent type of game. The decision-maker is matched with an informed expert and has a to take a sequence of decisions of unknown length. The decision-maker may verify the report of the agent but obtains no more feedback until the sequence ends. Finally, she cannot decide to substitute the expert. This feature is the main departure from previous literature (e.g. Ely and Valimaki (2003)) and can be interpreted in two different ways. First, it may be that the agent has been hired prior to this sequence of decisions and has secure tenure[7]. Alternatively the agent works on his own behalf and, therefore, he cannot be fired[8]. Finally, my model does also contribute to the (broad) literature on auditing games. The novelty from this perspective is the fact that the audited player can build a reputation of being "honest" and, therefore, affect subsequent auditing. It highlights the role of the time-horizon for the auditor under strong incentives.

This chapter contributes to the literature on reputation in dynamic games of asymmetric information. In particular, it contributes to a small but important literature that shows the limits of reputation as a way to restrain the behavior of patient players. The first paper to identify a shortcoming of the reputation effect is Morris (2001) who presents a model where an advisor with long-run concerns and no intrinsic bias against any alternative has an incentive to misreport his information in order to look "good". This incentive results in an equilibrium without any useful information provision. In a similar fashion, Ely and Valimaki (2003) study the sequential equilibrium of a dynamic game where the interaction between reputation and the interest of the agent to maintain his employment status, leads to no trade in equilibrium. The no-trade result is originated, as opposed to my paper, by the fact that the principals are short-lived and there is an information externality between them. Each of the principals would like to know whether she is facing a good or a bad agent but no one has incentives to incur in short-run losses in order to learn it. This externality is internalized by a long-lived principal who is able to commit and delegates decision-rights to the agent.

[7]For instance a newspaper hires a journalist to write a sequence of news pieces. The veracity of these is uncertain for the journal. Firing the worker is costly and may yield a reputation loss for the newspaper.

[8]Newspaper columnists, blogging economists or lobbyists offering advice on public policy are examples of this

Notice that in both Ely and Valimaki (2003) and this chapter the time horizon and patience of the uninformed principal play a key role in determining the equilibrium. In Ely and Valimaki (2003), information comes only after trade, so that for a myopic principal such information is of no use. In my paper, however, monitoring yields information ex-ante and a myopic principal may therefore have incentives to experiment. What is more, I show that, under some conditions, she has more incentives to experiment when she is myopic than when she is patient.

Another paper that is close to this one is Sobel (1985). He analyzes a similar game but assumes that output is perfectly and immediately observed by the principal and that the agent is perfectly informed about the state of nature. He concentrates on the conditions for equilibria involving information revelation by the agent at the beginning of the game. At some point, however, the agent will lie and the principal will know it for sure. Once this happens the game ends. Therefore, the goals of his paper are very different from mine.

As mentioned above, I also contribute to a small literature on endogenous information acquisition in games with repeated interaction. Liu (2011) presents a model with a sequence of short-lived buyers who have access to a costly technology that allows them to observe a number of previous trades of the seller. The seller may be a commitment or a strategic type, and buyers acquire a finite number of observations in equilibrium. The model predicts a cyclical behavior in reputation, with strategic sellers creating reputation until no buyer can distinguish them from a commitment type and then ”milking it down”. In my model past trades are observable and information acquisition concerns current actions. Kandori and Obara (2005) develop a sort of Folk Theorem for games of costly, private monitoring. Players monitor (audit) other players' behavior randomly and during cooperative phases, players both cheat and audit at the same time. There is, however, no reputation effects since there is no asymmetric information about players' types. In any case, my paper is the first to show that dynamic incentives for information acquisition may fail to give lead to more effort, in the presence of reputation effects.

More broadly, my results are also related to a branch of the literature on mechanism design that highlights the benefits of ignorance and uncertainty in the provision of incentives. Holmstrom (1999) presents a signal-jamming model in which agents are rewarded depending on the market belief about their ability. Ability and effort are substitutes, and so there is an incentive to provide more effort whenever the market belief is less precise. The main difference is that in this model ignorance fosters incentives *through the expected learning* and the dynamic value of information is always positive, and, therefore, the longer the horizon the higher the incentives to acquire information.

The remaining of the paper is organized as follows. In Section 1.2 I introduce the (infinite horizon) model and analyzes the main trade-offs in the simple static case. At the end of this Section I present a discussion of the main assumptions. In Section1.5 I solve the game played by a long-run agent and a myopic principal. This is the benchmark for my main result. In Section 1.6 I analyze the game played by two patient players. In Section 1.7 I study the role of commitment for the principal. In Section 1.8, I apply the model to some of the environments already mentioned. Finally, in Section 1.9 I conclude the paper.

1.2. The Model

1.2.1. Environment. Consider a game played by two infinitely-lived players: a principal (she) and an agent (he). The principal has to make a decision every period $d_t \in D = \{a, b\}$. In order to fix ideas, you may think of this decision as whether to invest in a given project. This decision generates a flow payoff that depends on the realization of the state of the nature $\omega_t \in \Omega = \{A, B\}$. I assume, for simplicity, that the ex-ante probability of both events is the same, so that $\Pr(\omega_t = A) = \Pr(\omega_t = B) = \frac{1}{2}$. She does not observe the realization of ω_t but she has access to a costless report by an agent (he) who has received has received a costless signal s_t correlated with the realization of ω_t. I assume a simple, binary structure for the signal, so that $s_t \in \{S, F\}$, where $s = S$ represents the case where the agent has acquired evidence in favor of alternative A and $s = F$ the case in which he has not. There is no hard evidence in favor of alternative B[9]. I denote by q be the posterior that the agent holds if he has obtained evidence and $1 - q$ the posterior in case he has not, so that q is the precision of the signal. Finally notice that neither the agent nor the principal do observe the realization of ω_t.

The principal can monitor the agent. By paying a fixed cost c, the principal verifies his report, requiring the agent to disclose the evidence he may have. Following with the banking example, the loan officer may decide to audit the accounts presented by the firm. Notice that given the asymmetric structure, and the fact that the agent may be biased in favor of A, the principal will never monitor after a report recommending B. This simplifies the analysis without affecting the main trade-offs.

The principal is a self-interested, forward-looking, rational player. She discounts the future by $\delta \in [0, 1)$ and maximize their expected

[9] Under the interpretation of B as no investment and A as investment. While it is possible to show that a project is profitable, it is very difficult to show that there does not exist a profitable project.

discounted stream of utilities. The principal's per-period utility function is

$$v_p(d, \omega, s) = u(d, \omega) - cm$$

where $u_p(d, \omega) = 1$ if $d = a$ and $\omega = A$ or if $d = b$ and $\omega = B$, and $u_p(d, \omega) = 0$ otherwise. Finally cm is the cost of monitoring. If the principal verifies the signal of the agent $m = 1$, and if he does not $m = 0$.

There are two types of agents. The *bad* agent does not share the same preferences over actions than the principal. In particular, the agent may enjoy a private benefit $\lambda > 0$ if action A is implemented. I will assume throughout that $\lambda > 2q - 1$ so that he always prefers action A independently of the state. On the other hand, if the agent reports the truth in every period, we say that the agent is *good*[10]. The agent's type is also his private information and is set fixed through time. The principal holds an initial prior belief μ_0 that the agent is good. Finally, as in Sobel (1985) assume that the "importance of the decision" for the agent varies from period to period, so let $x_t \in X$ represent this weight. Assume that $X = [x_l, x_h]$ is distributed according to $F(x)$, i.i.d through time, and is private information for the agent. As a normalization let $\mathbb{E}[x_t] = 1$. You may think x as a device to purify mixed strategies for the agent. Thus, if the agent is bad, his flow utility is

$$v_A(d, \omega) = [u(d, \omega) + \lambda \mathbf{1}_{d=A}]\, x$$

The following assumption simplify the problem.

ASSUMPTION 1.1. $x_h(1 - 2q + \lambda) < x_l q$

Assumption 1 guarantees that no matter how low is the realization of x_t the agent never prefers to recommend decision B when the signal he has received is S. It is sufficient to simplify both the problem of the principal and the agent without distorting the main intuitions.

The timing of the stage game is as follows. At the beginning of the period the principal holds a belief $\mu_t \in [0, 1]$ that the agent is good. The state of nature and the signals are drawn according to their distributions. The agent sends a report $r_t \in R$, the set of admissible messages[11], and the principal decides whether to verify the report or

[10] We assume that the *good* agent is a commitment type. It is straightforward to construct a payoff type choosing the same strategy. The complication is with respect to beliefs off-the-equilibrium path, since commitment types never deviate and payoff types may deviate. For a discussion see Mailath and Samuelson (2006), Chapter 15

[11] Since the *good* type only receives information about the state, if the *bad* type pretends to be a *good*, she may send at most two messages in equilibrium. Thus, without loss of generality R contains two elements. Hence, I shall assume that $R = \Omega$

not ($m_t \in \{0,1\}$). Finally, the principal chooses $d_t \in \Omega$ and the game moves to the following period.

1.2.2. Strategies and Equilibrium Concept. The main purpose of this paper is to analyze the (Stationary) Markov-Perfect Equilibrium of the infinite-horizon game. This equilibrium concept captures the idea that players have no commitment power and condition their strategies only on payoff-relevant information. Define a (private) history for the agent at the beginning of period t,

$$h_t^a = \{(s_k, r_k, x_k), (m_k, d_k)\}_{k<t}.$$

Let $\mathcal{H}_t$ be the space of such histories. The *bad* agent observes the realization of his private information (s_t, x_t) and decides which report to send. Therefore, a pure strategy for the agent is

$$r_t : h_t^a \times \{S, \emptyset\} \times X \to \Omega.$$

Recall that the good agent always tells the truth. A public history when the principal has to decide whether to monitor is

$$h_t^p = \{(r_k), (m_{k-1}, d_{k-1})\}_{k\leq t}. \tag{1.1}$$

Define $\mathcal{H}_t^p$ to be the space of such histories. A pure monitoring strategy for the principal

$$m : h_t^p \to \Omega$$

Stationary Markov strategies map payoff-relevant component of histories into actions with the constrain that for every two different histories h_t, h_r' generating the same posterior probability that the agent is good, the equilibrium play is the same. Thus, Markov strategies are measurable with respect to the posterior probability after every history. In what follows I describe these Markov strategies and present the concept of Markov Perfect Equilibrium.

The principal has to choose his actions using only payoff-relevant information. Given that both states are equally likely his choice of d_t is trivial. She will choose $d_t = a$ if and only if her posterior belief that the state $\omega = A$ exceeds $\frac{1}{2}$. For this reason, in the following, I will consider exclusively on the principal's choice of whether to monitor or not after a report A. Let $\sigma(\mu) \in [0,1]$ be the monitoring mixed strategy for a principal who holds a belief μ that the agent is good and receives a report A.

The bad agent, on the other hand, will misreport his information whenever it is in his best interest to do so. Thus, let $r(x;\mu) \in \{0,1\}$ be the probability that a bad agent, who has not received information and weights this period with x, and is believed to be good with probability μ by the principal sends report A.

Denote $V^k : [0,1] \to \mathbb{R}_+$ be the value function of the principal who has received a report $k \in \{A, B\}$ as a function of the current belief. Similarly, denote $U : [0,1] \times \{S, F\} \times X \to \mathbb{R}_+$ be the value function of the (bad) agent who has obtained a signal s_t and weights the current period with value x_t and is believed to be good with probability μ_t.

Let $\bar{F}(y) = 1 - F(y)$. Define $\phi_k(\mu; r)$ for $k \in \{A, B\}$ to be the posterior probability that the agent is good conditional on a report k if the prior was μ and the bad agent uses strategy r. With some abuse of notation I will write

$$\phi_k(\mu) = \int \phi_k(\mu; r(x; \mu))dF(x)$$

to be the equilibrium posterior probability from the point of view of the principal. Let $\pi(\mu; r)$ be the expected payoff for the principal after a recommendation to choose A of an agent believed to be good with probability μ who uses strategy r, if the principal is not monitoring.

DEFINITION 1.1. A Markov-Perfect Equilibrium is a pair of Markov strategies (σ^*, r^*) and a belief μ satisfying;

i) $\sigma^* > 0$, iff

$$q - c + \delta\mathbb{E}_{x,r}\left[V\left(\mu'\right)\right] \geq \pi\left(\mu; r\right) + \delta V\left(\phi_A(\mu)\right) \tag{1.2}$$

ii) $r = A$ if and only if

$$1 - q + \lambda + \delta\mathbb{E}_{\sigma,x}\left[U\left(\mu'\right)\right] \geq q + \delta\mathbb{E}\left[U\left(\phi_B(\mu)\right)\right] \tag{1.3}$$

iii) μ' is computed from μ using Bayes Rule[12]

1.3. Static Game

1.3.1. Known Type. It is useful to begin the analysis looking at the case in which the principal knows at the outset the type of the agent[13]. In this case, if the agent is good, there is no monitoring and the relationship yields an expected discounted payoff $\bar{V} = q$ for both players. On the other hand, if the agent is biased, there is a continuum of equilibria. Let p^* be the probability with which the agent reports A without supporting evidence in every period. This is the perceived probability of the principal, but the agent may use x as a randomization device. In every equilibrium, the principal will only monitor with probability 1 after an $A-$report whenever

$$c \leq p^*\left[q - \frac{1}{2}\right]$$

and never monitor otherwise. On the other hand, for the agent it is a (weakly) dominant strategy to report A after failing to obtain any

[12]From the point of view of the principal, every report of the agent has positive probability as long as $\mu \in (0, 1)$. If μ is degenerate, it will remain constant in the future. Therefore, off-the-equilibrium path beliefs are irrelevant.

[13]Formally, the type of the agent does also include the collection of realizations of the uncertainty (x, s). However, for lack of a better word, I will devote the word type to refer to his preference profile.

signal since

$$1 - q + \lambda > q.$$

Let p_0 satisfy

$$c = \frac{1}{2} p_0 \left[2q - 1\right]$$

ASSUMPTION 1.2. $p_0 < 1$

The assumption guarantees that monitoring occurs in equilibrium.

A particular class of equilibria that will turn out to be useful are threshold equilibria. For every $x^* \in \left[x_l, \bar{F}^{-1}(p_0)\right]$ there is an equilibrium in which the agent reports A if $x \geq x^*$ and the principal monitors if A is reported. Each of these equilibria yield the same payoff for the agent, but not for the principal. To see this, define V_{x^*} as the value for the principal in an equilibrium where the agent uses x^* as his threshold strategy.

$$V_{x^*} = q - \frac{1}{2} c \left[\bar{F}(x^*) + 1\right].$$

For the remaining of the paper, I shall consider only undominated strategies (see discussion below). It is obvious that reporting B is (weakly) dominated for the agent, in the symmetric information game, so that I set $x^* = x_l$. Thus, the value for the principal is

$$\underline{\mathrm{V}} = q - c$$

1.3.2. Unknown Type. We consider now the case in which the principal is uncertain about the type of the agent. Suppose that the probability she attaches to the event that the agent is good is μ. Then, she will monitor the agent if and only if $\mu \leq \bar{\mu}$, where $\bar{\mu}$ is defined in

$$c = \frac{1}{2}(1 - \bar{\mu}) \left[2q - 1\right]$$

and so the principal gets a value

$$V(\mu) = \mu V_g(\mu) + (1 - \mu) V_b(\mu)$$

with $V_g(\mu) = \frac{1}{2}q + \frac{1}{2}(q - c) > V_b(\mu) = q - c$ if $\mu \leq \bar{\mu}$ and $V_g(\mu) = q > V_b(\mu) = \frac{1}{2}$ if $\mu > \bar{\mu}$

Finally, the value for the principal of learning the type of the agent, in this static framework is

$$\begin{aligned} \Pi &= \mu(\bar{V} - V_g(\mu)) + (1 - \mu)(\underline{\mathrm{V}} - V_b(\mu)) \\ &= \left\{ \begin{array}{c} \frac{1}{2}\mu c \text{ if } \mu \leq \bar{\mu} \\ (1 - \mu)(q - \frac{1}{2}) \text{ if } \mu > \bar{\mu} \end{array} \right. \geq 0 \end{aligned}$$

Obviously in this simple game, the principal is willing to pay to know the type of the agent. In the rest of the paper we show that this incentives need not be present in the infinite-horizon version of the model, since information *does* change the continuation equilibrium.

1.4. Discussion

Although in a very stylized way, this model is able to capture the main features of the strategic interactions presented above. Think, for instance, in the game played by the bank and an entrepreneur. The entrepreneur may come, every period, with a potentially profitable project. He may also ”save for better times” if there is no profitable project at hand. The bank would like to finance only profitable projects, but ascertain whether a project is profitable is costly. The bank would like to use the informational content of the decision of the entrepreneur to offer a project, in order to save resources from costly monitoring. This is possible, if the entrepreneur builds a reputation of being ”trustworthy”, by behaving as a ”good” type.

As mentioned in the Introduction, the main feature of the environment is that the agent is paired with the principal and there is no exit option. This assumption may seem restrictive, but I feel that it captures quite well the environment described. First, in family relations exit is usually infeasible. Moreover, in those relations that allow for exit, it is usually very costly since the agent may be very hard to substitute. Our results would not change if verifiable evidence of misbehavior is required for dismissing the agent, and would be qualitatively similar if the principal dismisses the agent whenever his belief on his type falls below some threshold, since the good agent is committed to telling the truth and not strategic.

I also assume, in this particular version of the model, that players do not learn their payoffs as they play. All results will go through as long as output is not perfectly observed (or $q < 1$) since information about payoffs is received after decisions are made. This information does not change the updating of interim beliefs, and, therefore, is irrelevant when analyzing the strategic interaction between the agent and the principal. The limiting case in which output is immediately and perfectly observed and $q = 1$ is analyzed in Sobel (1985)[14].

Finally, I assume that stage-game payoffs are independent of the current belief that the market has. This is a simplification, and it is unlikely to hold in real-world examples. It requires, for instance, that, *a priori*, both alternatives are equally likely and that the importance of the period x_t is distributed independently of the reputation of the agent. Removing any of these restrictions will not affect the results, as long as the stage-game payoff of the agent is monotonically increasing in the belief that the principal holds about his type. This is satisfied in all the examples introduced. For instance, in the credit market example, interest rates will be decreasing in the belief that

[14]Details of this extension are presented in a Supplemental Appendix to Garcia-Gonzalez (2012).

the lender has about the agent's type and, therefore, the payoff of the agent would be increasing in this belief.

1.4.1. Solution Concept. In the remaining of the paper I focus on undominated equilibria. In particular, I rule out weakly dominated strategies in subgames that involve perfect information and monitoring with probability one. Notice that these equilibria rely on dominated strategies for the agent but are weakly preferred for the principal. Alternatively, I could have assumed that the monitoring technology fails to give any information with some positive, albeit small, probability. Undominated equilibria will be the only equilibria surviving in such perturbed model. In this sense, the equilibria I look at is the limit with respect to a class of imperfect monitoring technologies, while the rest are not.

I also restrict attention to Markov strategies. I believe that Markov strategies are the natural way to introduce lack of commitment and impose sequential rationality in infinitely-lived relations. Notice that a strategy is Markov if it is only measurable with respect to payoff-relevant information in every period in time, so that for every two histories originating the same beliefs Markovian strategies specify the same distribution over actions. In particular, the principal is unable to commit to future punishments conditional on deviations that do not trigger a change in his beliefs about the type of the agent. This kind of punishments require players to coordinate either through correlating devices or their history, which may be difficult in many situations of interest. Notice that my aim in this paper is to compare the situation where both the agent and the principal are long-run players with one in which the principal is short-run. Equilibria involving future punishments and rewards renders the comparison between the two environments difficult.

Markov Perfect Equilibria are formed by simple strategies. This simple structure is well-suited for studying interactions within organizations, where decision-making follows standardized protocols and routines requiring the use of limited information. Organizations tradeoff the benefits of more accurate decision-making versus the costs of information transmission and acquisition (Arrow (1974)). Therefore, strategies requiring infinite recall and the use of unbounded information are not useful for understanding the equilibrium behavior of real-world organizations.

1.4.2. Monitoring Technology. I assume that monitoring is both public and perfect. This means that the agent knows for certain whether the principal has verified his report and the result of such verification. This assumption is justified in many of the environments described above, e.g. relationship finance, where the agent is asked to disclose his information.

The assumption is made mainly for the sake of simplicity. First, if monitoring were private but perfect, the state would not only include the current belief of the principal but also, the second-order belief of the agent. In general, this second-order belief is distribution over the set of possible beliefs that the principal may have (depending on her past actions). In my simple model, however, the principal is not willing to distort her decision in order to conceal her information, since the very reason to monitor is to make better decisions. Therefore, the second-order belief will be degenerate and no additional insights will be obtained.

Second, one could assume that the monitoring technology is imperfect, in the sense, that it yields no information with some positive probability. This is the technology used in Diamond (1991) among others. My results are robust to this extension and, moreover, our equilibrium selection criterion picks the limit equilibria of the class of games indexed by these technologies where the probability of failure goes to zero. But allowing monitoring to be both imperfect and private would affect the results non-trivially and is left for future research.

1.5. Myopic Principal

I will present now the results for the benchmark model where the principal is myopic while the agent is patient and discount the future with a factor δ. The strategic agent maximizes, then, the present expected discounted sum of his stream of payoffs. After the agent observes his signal s_t and the importance of the period x_t, his value function is

$$U(\mu; x_t, s_t) = \max_r x\mathbb{E}_\sigma \left[v_A(d, s) \mid r\right] + \delta\mathbb{E}_{x,\sigma}\left(U(\mu) \mid r\right)$$

with

$$U(\mu) = \frac{1}{2}\int \left[U(\mu; x, a) + U(\mu; x, b)\right] dF(x)$$

where the agent takes expectations over the (mixed) strategy of the principal and the future realizations of the uncertainty. The principal will never monitor with probability 1 since in that case the agent will not be lying with positive probability. On the other hand, if the agent plays a threshold strategy x[15], the principal will randomize, and monitor with positive probability only if

$$c = (1-\mu)\bar{F}(x^*)(2q-1)$$

This determines x as part of the equilibrium for those μ such that $\sigma(\mu) > 0$. Call $x_s(\mu)$ the solution to the equation (in case it exists).

[15] In the long-run case, every equilibrium is outcome equivalent to a threshold equilibrium.

We need to find $\sigma \in (0,1)$ to make $x_s(\mu)$ optimal

$$[x_s(\mu)(1-q+\lambda)+U(\phi_A(\mu))](1-\sigma)+\sigma[qx_s(\mu)+U(0)] = qx_s(\mu)+U(\phi_B(\mu))$$

The following Proposition summarizes this equilibrium [16].

PROPOSITION 1.2. *In any equilibrium of the game with one long-run player, the principal monitors with positive probability if and only if μ is small enough, the agent lies with positive probability for every μ. The (strategic) agent is discovered in finite time.*

The typical play is as follows. If the principal is confident that the agent is good, she does not monitor and the agent fully chooses the future path of posterior beliefs. In this case, he restraints his behavior because if the belief goes down, the principal will monitor and thus his influence in the decision will be lower. If the agent is bad, beliefs will follow a stochastic path with negative drift and, at some point, the principal starts to monitor with positive probability. In this case, the agent's payoff from suggesting option A is decreased but he still lies with positive probability, until he is discovered. Importantly, the agent does always restraint his behavior as compared with the one-shot interaction case if $\mu \in (0,1)$, in the sense that he cooperates and tells the truth with positive probability. This increases the equilibrium payoff of the principal. Finally, notice that this allows the principal to reduce his monitoring (if the agent chooses B the principal does not bear the cost of monitoring). This is the key observation for what follows.

1.6. Patient Principal

The problem for the agent is similar than the one presented above, given that he takes σ as given. The problem of the principal changes in two dimensions. First, the principal now has an incentive to learn the information of the agent to make better decisions and thus increases his sustained payoff. But second, he faces a dynamic inconsistency problem that will lead him to reduce the speed of learning and thus, the reputational concern of the agent. Characterizing this second force is the main aim of this paper.

Whenever the agent chooses action A, the principal can choose between verifying his signal or not. If he does, his future discounted expected payoff is

$$V_{A,1}(\mu) = q - c + \delta\left[(1-\mu)\bar{F}(x(\mu))V(0) + (\mu + F(x(\mu)(1-\mu))V(\mu)\right]$$

while if he does not verify

$$V_{A,0}(\mu) = \pi_\mu + \delta V\left(\phi_A(\mu)\right)$$

[16]If the Principal is short-lived Existence of Stationary MPE can be shown using slight modifications of the results in Bar-Isaac (2003)

Notice that the principal randomizes if and only if

$$V_{A,1}(\mu) = V_{A,0}(\mu)$$

In Appendix A, I show that both V and U as well as the policy functions are monotone. Thus, minor modifications of well-known arguments guarantee equilibrium existence. I am now ready to state the main result of this Chapter. Namely, that in any MPE the patient principal will reduce his monitoring intensity at some beliefs.

PROPOSITION 1.3. *In any equilibrium of the game in which both the principal and the agent are long-run players, there exists some $\mu_l > 0$, such that if $\mu_t < \mu_l$, $x^s(\mu_t) > x^l(\mu_t)$ and $\sigma^s(\mu_t) > \sigma^l(\mu_t)$*

To see this result, consider the value function as a weighted average of the two conditional value functions $V(\mu \mid t)$ for $t \in \{g, b\}$. It is clear that $V(\mu) = \mu V(\mu \mid g) + (1-\mu)V(\mu \mid b)$ where μ defines the equilibrium strategies and the type determines the payoffs. Notice that the martingale property of the beliefs is such that the dynamic value of information is pinned down by a weighted average of the change in those two value functions as information arrives. In particular $V(\mu \mid g)$ is decreasing in μ since monitoring is redundant in this case, and takes the value $\frac{q}{1-\delta}$ for all μ such that $\sigma(\mu) = 0$. However, $V(\mu \mid b)$ is not maximized at $\mu = 0$ since the reputation effect is absent. In particular, $V(\mu \mid b)$ is bounded away from its supremum at $\mu = 0$. Intuitively, if the principal is indifferent of whether to monitor or not, conditional on receiving a report recommending action A, the informativeness of such report must be bounded away from zero. Therefore, as the belief goes to zero, the probability of a lie must remain bounded away from 1. This implies that the value for the principal of maintaining uncertainty on the agent's type is strictly positive for every μ and the result follows.

The following proposition shows that this may increase the value for the agent for every $\mu \in (0,1)$

PROPOSITION 1.4. *There exists some $c_1 > 0$ such that if $c \geq c_1$ then the agent lies with higher probability when matched with a long-run principal ($x^s(\mu) \geq x^l(\mu)$) and gets a higher payoff, $U^l(\mu) \geq U^s(\mu)$, for every $\mu \in (0,1)$*

As mentioned above, these are the main results of the paper and show how a principal facing an agent who is likely to have conflicting interests with himself, faces a trade-off when deciding whether to investigate him. Absent any dynamic concerns, she may find it optimal to monitor him closely, in order to avoid being cheated, but whenever she has a long-term concern, she will reduce her monitoring intensity and try to free-ride on the dynamic concern of the agent. This will in its turn lead to a lower incentive for the agent to restrain his behavior and thus lower payoffs for the principal.

1.7. The Role of Commitment

In this Section I explore commitment or the principal. I will first allow the principal to commit at the beginning of the period to a monitoring intensity. This implies that the principal need not be indifferent whether to monitor or not conditional on a report recommending A and that she internalizes the effect of a higher monitoring intensity in the behavior of the principal. In short, the principal becomes the Stackleberg leader. Her problem becomes

$$V^C(\mu) = \max_{\sigma\in[0,1]} p_\mu(\sigma)\, q + (1-p_\mu(\sigma))\, [\sigma(q-c) + (1-\sigma)\pi_\mu(\sigma)] + \delta \mathbb{E}\, V^C(\phi(\mu))$$

where $p_\mu(\sigma)$ is the expected probability of a report B when the belief is μ and the principal committed to a strategy σ. Notice that $V^C(\phi(\mu))$ may be not differentiable. However, it is easy to see that the maximization problem is globally concave[17]. The following proposition shows that the solution to this problem $\sigma^C(\mu) \geq \sigma^l(\mu)$ for every μ.

PROPOSITION 1.5. *If the principal has long-run concerns and is able to commit to a monitoring strategy before the agent reports, she will monitor with higher probability than in the case in which she cannot commit.*

The proof uses the fact that at $\sigma^l(\mu)$ the principal is ex-post indifferent between monitoring or not, but ex-ante has an incentive to increase monitor to lead to a higher accuracy of the report and save on monitoring costs. Therefore, if $\sigma^l(\mu) > 0$, $\sigma^C(\mu) > \sigma^l(\mu)$. Finally notice that this implies that whenever $\sigma^S(\mu) \geq \sigma^l(\mu)$ the principal may find it useful to delegate monitoring to a third party and give him short-term incentives.

The preceding discussion suggests that there is a role to commitment. But, can a principal who fully commits to a monitoring intensity get his first-best payoff? Ely and Valimaki (2003) show that in an environment similar to ours, if the principal is a long-run player and can choose a (stochastic) participation-rule as a function of the entire history, social first-best is attainable (under the average discounted expected payoff criterion). The idea of their mechanism is to ”promote” the agent after every period with positive probability, so that monitoring stops altogether. The probability is chosen to separate types, so that a *bad* type is not willing to incur in the cost of mimicking the good type and prefers to deviate in the first period. For this result, it is required that deviating today yields high payoffs for the *bad* type. In particular, the decision-rule today must follow his advice -although this advice is also known to be wrong.

[17]Informally, increasing σ reduces both the probability of a report A but also the difference in continuation values (less informative signals.)

In my model, unlike in Ely and Valimaki (2003), however, the decision is made by the principal. This implies that if the agent expects the principal to verify his report with high enough probability, the agent will pretend he is a good type and tell the truth. Hence, the principal will be unable to tell them apart fast enough and fail to get her first-best payoff. To understand the result, let $V^i : \mathcal{H} \to \mathbb{R}$ be the value that the principal can get after a given history $h \in \mathcal{H}$ where she to play against a type $i \in \{g, b\}$, and let $\mu(h)$ be the implied belief of the principal.

PROPOSITION 1.6. *For every* $\mu(h) \in (0,1)$*, if* $V^g(h) = \frac{q}{1-\delta}$ *then* $V^b(h) < \frac{q-c}{1-\delta}$.

The argument is simple. Since the bad agent may pretend to be a good type, in order to give incentives for revelation the principal must give him higher continuation payoff. But absent any commitment to follow the advice, the principal can only give higher payoffs by reducing her monitoring intensity. Lower monitoring implies delay in information and more cheating, so that her payoff is bounded away from first best. On the other hand, if she were to monitor more, her payoff on the good agent will be lower since monitoring is inefficient in such a case.

As it is clear from the proof the incentive compatibility constraint of the agent is binding and this implies that revelation will be slow enough for the principal to obtain his first best value. If the principal could commit also to delegate the decision to the agent, this would soften this constraint and increase the value of the principal. Thus, an important insight of this paper is that delegating monitoring and decision-making may be superior to delegating monitoring and relying on communication.

1.8. Applications

Organizational Design: As mentioned in the Introduction, this paper offers novel insights about the use of different institutions in many organizations. First, we provide a new rationale for delegated monitoring. In the framework presented her, monitoring is not only required to provide incentives but also to acquire information about the type of the agent. The value of such information is not always positive, since it reduces future incentives. Therefore, delegating monitoring to a third party who does not internalize this negative value of information may increase monitoring and provide better incentives now.

Second, we provide an intuitive explanation for the pervasive use of short-term incentives in many organizations where information is revealed over time and cooperation is required for efficiency (Hertzberg et al. (2010)). Agents with long-run concerns may have low incentives to monitor agents with whom they interact, since such monitoring may bring bad news and decreases the continuation payoff.

Third, in my model, even if monitoring is delegated to a third party, the principal may want to delegate the decision to the agent. Under delegation, the *bad* agent is willing to reveal his type more quickly and first best may be approximated. If the agent is to report to the decision maker, however, the agent will conceal his information and preclude first best. Hence, I give an additional rationale for the widespread use of delegation, even when information is potentially verifiable.

Relationship Finance: One of the main applications of the model presented here is the banking industry. In particular, my model is very similar to that in Diamond (1991), who is the first to study the interplay between monitoring and reputation but under the assumption that the monitor is short-lived. Among other things, he discovers a "paradox of monitoring" whereby cheap and effective monitoring fails to create incentives since lenders face a lack of commitment to cut defaulters off the credit market. I extend the analysis to incorporate long-run motives for the monitor. I show that, if the credit score of the firm is sufficiently low and the bank has a lack of commitment, dynamic interactions may not be able to solve such problems[18].

More generally, my model sheds light on the use of relationship finance versus arm's-length or directly placed financing. Relationship finance is widely understood as a way to overcome informational asymmetries between borrowers and lenders. Banks accumulate information about their borrowers over time, both through communication and monitoring. The value of communication depends on the incentives of the incentives of the borrower to maintain a reputation and the threat of monitoring. Communication saves on information acquisition costs and, therefore, is valuable for the bank. The model predicts that banks starting long-term relationships will use less external or formal monitoring and rely more on communication and soft information than borrowers engaging in arm's-length financing. This is consistent with empirical evidence presented in Kano et al. (2011) who show that firms benefit most from bank-borrower relationships when they do not have audited financial statements. Similarly, Blackwell and Winters (1997) finds that banks less frequently monitor firms with whom they have closer relationships. Even more,

[18]There is one additional difference between both frameworks. He carefully models the credit market and, therefore, allows interest rates to be endogenous. Interest rates will be monotonic in the belief that the market has about the agent, so that the per-period payoff of the agent will be endogenously increasing in this belief. Adding this monotonicity would not change the results of my model, but in the interest of simplicity I ignored it.

many studies have found that banks lending to insolvent firms are more likely to execute their guarantees or liquidate assets if they do not have a long-run relationship with the borrower.

The existing literature has relied on two different theoretical frameworks to understand this evidence. First, banks may suffer a *soft-budget constraint problem* so that they are unable to commit ex-ante not to refinance inefficient projects since the initial investment is sunk Dewatripont and Maskin (1995). This lack of commitment, together with an agency problem, is most likely to be present in bank-oriented economies where banks perform relationship financing and have more ”at stake”. Even so, these refinancing decisions are sequentially efficient, in the sense that the expected value of those projects at the time of the refinancing decision exceed their liquidation value. Therefore, this argument is insufficient to explain ”zombie-lending”, i.e. the concession of debts to inefficient firms at a subsidized rate, e.g. Caballero et al. (2008), since these rates are computed using all the available information at the time of the decision.

The second common explanation relies on an *efficiency-wage* type of argument. Banks may lower their monitoring effort because they provide rents to entrepreneurs (lower interest rates) and, therefore, alleviate the moral hazard problem. This hypothesis requires that the entrepreneur is easily substitutable for the bank, and that the bank has commitment to do so.

According to the hypothesis developed in my paper, however, the terms of refinancing offered by banks to long-term borrowers will be the less sensitive to the financial situation behavior than that of firms with whom they have short-run relationships, and more so, whenever these firms are in financial distress. Banks face the risk of losing the relationship they created with a firm if it goes under and this reduces the incentives they have to monitor their activities. This, in turn, decreases the incentives for firms to devote resources to efficient activities and, thus destroys value for the bank.

Nonetheless, it would be a mistake to conclude for this argument that relationship financing is not profitable for banks. As in Diamond (1991), monitoring generates valuable information about the type of the agent, which has a future value only in the case of relationship financing. The novelty of this model stems from the behavior of the agent, which is also affected by the current belief of the bank about his type. Changes in behavior induced by changes in beliefs, decrease the incentives to monitor for the principal and eventually overcomes the positive effect of information generation. This is the trade-off analyzed in the current paper.

1.9. Conclusions

In this chapter I have presented a simple dynamic framework to understand the interaction between an uninformed decision-maker and an informed agent who may be bad in favor of one of the alternatives. The decision-maker may actively monitor the agent but lacks commitment and is bound to suffer losses in case that she discovers that the agent is bad. In this environment, we show that dynamic concerns for the agent increase the payoff of the principal, but dynamic concerns for the principal may decrease her payoff, if her belief about the type of the agent is sufficiently pessimistic.

The key observation of the paper is as follows. In any equilibrium with positive value of reputation, the principal must be (at most) indifferent between auditing or not. Therefore, the agent will not lie with probability 1 in any equilibrium. This increases the average value for the principal as long as the agent has not been found to be bad with certainty. Monitoring more intensively increases the probability that this happens, and, therefore, reduces the future value of the principal. This implies that the dynamic value of monitoring becomes negative when the probability that the agent is bad is sufficiently high, and, therefore, the long-run principal monitors less intensively. This increases the value of the agent and gives him more incentives to cheat.

We have also shown that even if the principal is able to commit to a fully contingent plan of future monitoring intensity, he can not approximate her first-best payoff. This is because the principal retains the decision rights even if monitoring was delegated, and once she discovers the type of the agent, has no incentive to bias his decision. Therefore, the incentive for the agent to reveal his type is diminished and the agent has to monitor for a sufficiently long period of time that his payoff is bounded away from first-best.

Finally, we have presented a number of different environments where this insight seems to be present, and discussed the implications of our findings for organizational design, highlighting the benefits of delegation and short-term incentives and the shortcomings of communication. As for future research, it seems interesting to exploit other organizational arrangements that may allow the principal to commit and limit wrongdoing by the agent.

CHAPTER 2

Communication and Information Acquisition in Networks

'The single biggest problem in communication is the illusion that it has taken place'. George Bernard Shaw

2.1. Introduction

Economists have long recognized the acquisition and transmission of information between individuals as one of the key objectives of organizations (Arrow (1974).) Indeed, organizations take over the role of prices when these fail to accomplish their mission of aggregating disperse information and encouraging individuals to take the appropriate actions. Mimicking the role of prices in market transactions, organizational design should enable efficient information transmission within the organization and provide the right incentives to create and maintain information flows from outside. While these two elements have been separately studied in different papers[1], this chapter constitutes the first attempt to analyze their interrelation and their implications for organizational design. I argue that this link may explain some of the features of many real-world organizations.

A good example of an organization in which information transmission is important is the stock market. Most information is conveyed through prices, but it is also well-known that word-of-mouth communication and other networked activities are ubiquitous in those environments. Shiller and Pound (1986) shows that most trading decisions involved interpersonal communication, and very few agents make their own research. Similarly, Hong et al. (2005) finds strong correlation in the positions of traders based on the same city, controlling for the location of the assets. This evidence suggests a strong use of personal contacts in information acquisition. This has been neglected in the majority of papers studying financial markets, where

[1] Bergemann and Valimaki (2002) provides a general framework to study information acquisition. For studies in communication, see Crawford and Sobel (1982).

the information structure is a reduced-form stochastic process. In particular, no explicit distinction is made about the sources originating the signals[2].

This chapter highlights a bidirectional interaction between information acquisition and communication. First, smooth information transmission helps to disseminate relevant information and coordinate behavior, while reducing the duplication of efforts in information acquisition. But differences in the information available to different agents will hamper their (mutual) communication, since it introduces a wedge between the conditional expectations received after some signal is observed. Agents communicate their signals before obtaining all the relevant information and use interim beliefs which depend on the amount of information that they expect to receive. For instance, more informed agents rely *less* on every particular signal than less informed agents. This implies that their second-order beliefs will differ from their first order beliefs and information transmission will be noisy. Players have an incentive to become *conservative* when communicating information to less informed ones and *aggressive* when communicating to more informed ones.

To get a grasp of the implications of this trade-off, I study a standard beauty-contest type of game (Morris and Shin (2002)) where every agent must take a decision facing a trade-off between adaptation to global uncertainty and coordination with the rest of players. I allow them to choose the amount of information they acquire (Hellwig and Veldkamp (2009), Myatt and Wallace (2011)) and to report this information to their peers through a discrete (undirected) network. In the benchmark model, this communication takes the simple form of a round of messages in which every agent chooses a profile of reports to each of his peers conditional on the information he owns.

2.1.1. Applications. As already mentioned, the model captures the environment faced by communities of financial analysts. First, beauty contests are a useful, albeit simple, tool for trying to model financial markets and asset prices[3]. Second, information about financial assets is disperse and different traders may differ in the amount of information they have access to. Finally, information transmission among traders does not occur through a "market" but rather under

[2]Two recent studies constitute an exception to that rule, in that they address the impact of exogenous networked information structures on trading behavior. **?** studies a model with "behavioral" communication through simple topological structures and their results in trading. Ozsoylev and Walden (2011)) studies a rational-expectations equilibrium in which agents communicate truthfully through an exogenous random network. See Section 2.8.2 for a relation of our study with this literature.

[3] See, for instance, Allen et al. (2006).

bilateral, stable relations which we characterize using the network device.

Nevertheless, this model may well be applied to many different settings in which communication is strategic and unverifiable. For instance, it may be useful to understand information sharing between firms operating in similar markets where strategic complementarities are present (Raith (1996).) Pairwise communication is potentially less costly and more difficult to detect but may create problems in terms of credibility. Similarly, our framework may be applied to the study of complex organizations[4] whenever information is disperse and coordination is key for performance. In these organizations, decisions must be taken rapidly[5] and communication is informal. For instance, coordination, information acquisition and good communication are the key factors underlying the design of Intelligence Agencies (Garicano and Posner (2005)). In this model we argue that their interaction may indeed reduce performance.

2.1.2. Overview of the Results. In this chapter, I show that differences in the amount of information different agents have access to difficulties the communication among them and that more information acquisition generates an external effect which may be positive or negative. Our results show that truthful information transmission depends on both the network topology and the information acquisition technology. In particular, very centralized structures (formally, core-periphery networks) or very decentralized (regular networks) yield efficient communication, while hybrids typically don't. We also show that, in line with previous literature, information acquisition is monotonic in the centrality of the player if communication is truthful, but it may be not monotonic otherwise. The intuition for this result stems from the fact that the degree of substitutability of information between players is *endogenous* to the structure of communication. In this sense, information is a non-rival good if and only if the organization allows for truthful revelation.

Finally, I show that our qualitative results extend to an environment with more rounds of communication, as long as players leave the network once they take their actions. In particular, I identify network structures for which, independently of the number of rounds of communication, information cannot be truthfully revealed between two linked players because they will use it differently. In this sense,

[4]A recent literature in the economics of organizations, starting from Dessein and Santos (2006) has used similar specifications. See Calvó-Armengol et al. (2011) for a discussion of the relation with this literature and a deeper analysis of its implications to organizational design.

[5]See Section 5 for a brief analysis of the cost of delay in decision-making and its implications for communication.

the fact that our results rely on the use of interim beliefs does not imply that the assumption of one round of communication is crucial[6].

2.1.3. Related Literature. This chapter contributes to a couple of strands in the literature. First, there is a small but influential literature on communication in networked organizations started by Geanakoplos and Milgrom (1988)[7] (1991) and Radner (1993), within the realm of team theory. There are no strategic issues and the problem is simply to choose the optimal organization of workers to minimize time processing, due to bounded rationality. The typical finding of this literature is that hierarchical organizations are likely to be optimal for information transmission purposes. Adding strategic incentives to the transmission of information, we find that hierarchies are likely to be suboptimal since they yield a very unequal distribution of information and, therefore, weak incentives for truth-telling.

Second, there is a growing literature of game-theoretical views of networked organizations. Calvó-Armengol et al. (2011) study information acquisition and truthful and costly communication in networks. They show that information acquisition is increasing in centrality in a linear-quadratic model of network formation. However, this chapter is the first study addressing information acquisition and strategic communication jointly. Finally, a couple of recent contributions deal with strategic information transmission in networks. Hagenbach and Koessler (2010) analyze a game in which signals are strategic complements and agents differ on their preference relation over outcomes, but there is no information acquisition and the preference divergence is exogenous to the network structure. Galeotti et al. (2009) analyze a similar game, but their focus is on competing signals and analyze the effect of congestion and other network characteristics on the amount of information transmitted.

To conclude, two recent contributions analyze repeated communication in societies. Anderlini et al. (2012) consider an organization composed by one-period-lived agents who send reports to their successors regarding some underlying uncertainty. They show that the existence of an exogenous preference bias impedes common learning of the parameter of interest. Acemoglu et al. (2010) is somewhat closer to our spirit and they consider the case of large societies transmitting over time information relevant to the decision of whether to undertake or not a project. They highlight an strategic motive to lie to induce agents to transmit their information, but they concentrate mostly on truthful communication.

[6]The crucial assumption is that agents may take actions after each round of reports.

[7]See also Bolton and Dewatripont (1994)

2.2. Model

Consider a set N of agents. Let $2 \leq n < \infty$ be the cardinality of N. Every agent is concerned with the realization of some aggregate uncertainty θ. In the case of financial analysis, θ would be the fundamental value of an asset. I assume that θ follows a normal distribution, $N(0, \tau_\sigma^{-1})$. Agents, however, do not observe the realization of θ. They only receive a signal $x_i = \theta + \eta_i$, with η_i normal with zero mean and variance τ_i^{-1} where τ_i is the precision the signal held by agent i. Notice then that $\{x_i\}_{i \in N}$ are independent conditional on θ but may not be identically distributed, since we allow agents to choose some precision $\tau_i \in \mathbb{R}$, by paying some cost $c(\tau_i) \in \mathbb{R}_+$. For the moment, I only assume the cost function to be increasing and convex. Let $\tau_i(g)$ be the information acquisition strategy of agent i when playing on network g. Let $\tau = (\tau_1, \tau_2, .., \tau_n)$ be the profile of precision choices for each player, which is known at the end of the first period period[8]

Agents are linked through an undirected and discrete network g, so that i and j have a link if and only if $ij \in g$. A link is interpreted as the existence of a communication channel between two players. Alternatively, one can define the network by using a matrix G of zeroes and ones, such that the coefficient $G_{ij} = 1 \Leftrightarrow ij \in g$[9]. Define a *walk* from i to j as a collection $\{k_1, k_2, .., k_m\}$ such that $k_1 = i$, $k_m = j$ and $k_l k_{l-1} \in g$ for all $l = 2, 3, .., m$. A *path* is the shortest walk between two players, i, j, so that we write $p(i,j)$. Let $|p(i,j)|$ be its length. If such a path does not exist $|p(i,j)| = \infty$. A network is *connected* if and only if for all i, $sup_j \, |p(i,j)| < \infty$. A network is *minimally connected* if for every $i, j \in N$, there exists one and only one path linking them.

A *component* $g^s \subset g$ is a subnetwork of g such that the nodes of g^s, N^s, is a set of players satisfying for all $i, j \in N^s$, $p(i,j) < \infty$ and for all $k \in N \setminus N^s$ we have that $p(i,k) = \infty$. Let n^s be its cardinality.

Most of this chapter is concerned with one round of simultaneous communication among linked players, contingent on their private signals. Denote by $N_i(g) = \{j \in N : ij \in g\}$ the set of neighbors of agent i, so that every agent submits a message profile $m_i = \{m_{ij}\}_{j \in N_i(g)}$ and learns a message profile $m^i = \{m_{ji}\}_{j \in N_i(g)}$. For simplicity, I shall assume that $m_{ij} \in \mathbb{R}$, so that a reporting (pure) strategy for agent i is a mapping $m^i : \mathbb{R} \to \mathbb{R}^{|N_i(g)|-1}$.I denote with $N_{ij}(g) = N_i(g) \cap N_j(g)$ the set of common neighbors of i and j, and $N_{i-j} = N_i(g) \setminus N_{ij}(g)$ the set of neighbors of i who cannot communicate with j. Finally, let

[8]This assumption may be relaxed by imposing that agents get to know a noisy signal of the precision of other agents. For instance, one could allow the technology of information acquisition to be random and only the cost would be observed.

[9]For the most part we use the set-theoretic definition, since we find it more intuitive. However, some definitions are better presented in matrix form.

$N_i^*(g) = N_i(g) \cup \{i\}$ be the neighborhood of agent i augmented to himself[10].

To allow for mixed strategies in reporting let $\mu_{ij}(y \mid x)$ is the probability of sending report $m_{ij} = y$ conditional on signal x. I shall require that $\int_{-\infty}^{\infty} \mu_{ij}(y \mid x)dy = 1$. Let $\hat{m}_{ij}(x) = \{y : \mu_{ij}(y \mid x) > 0\}$ and $\hat{m}_{ij}^{-1}(y) = \{x : \mu_{ij}(y \mid x) > 0\}$.

In the last stage, players must take an action $a_i \in \mathbb{R}$, conditional on all the information available to maximize

$$U = -(a_i - \theta)^2 - \frac{1}{n-1}\sum_{j \neq i}(a_i - a_j)^2 - c(\tau_i) \tag{2.1}$$

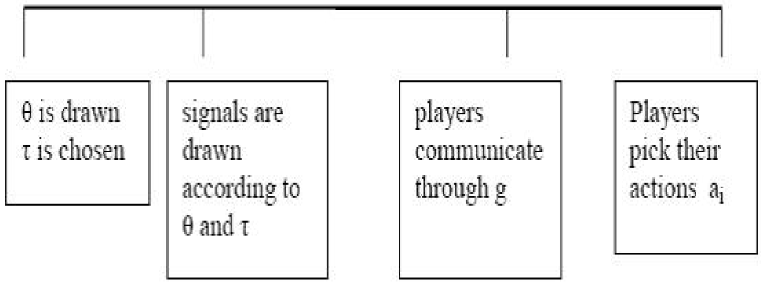
According to (2.1) every agent wants to match a weighted average of the underlying uncertainty and the actions of other players. However, it also implies that they would be trying to lead other agents to take the same action as they expect to take themselves when making their report. Let

$$a_i : \mathbb{R} \times \mathbb{R}^{|N_i(g)|-1} \to \mathbb{R}$$

be the strategy of agent i contingent on her information. We assume that transfers contingent on reports or actions are not possible in this environment. Transfers would be difficult for enforce since information is not verifiable. More importantly, once information has been acquired, players would like to communicate it at any cost -so that prices for reports may be negative (people are paid to listen.) Because of this, the market is unable to encourage research by pricing it in the absence of further commitments. Finally, the price of research of every agent should depend on both his position in the network and the effort of every other player, requiring very complicated schemes. As a matter of fact, contracts for information (e.g. consulting or advising services) require that providers have no intrinsic interest in the final decision. For instance, financial analysts are not allowed to make priced reports on assets on which they have interest.

To end this section, I summarize the timing of events, as displayed in Figure 1. In the first stage, nature draws a state of the world and every agent chooses some information acquisition τ. Signals are then drawn conditional on the state of the world according to the chosen distributions. In the second stage, every player communicates to her peers through an undirected network g. Finally, conditional on all her information, every player chooses an action a_i and payoffs are realized.

[10] We extend analogously the remaining concepts, so that, for instance, $N_{ij}^* = N_j^* \cap N_i^*$



2.3. Truthful Revelation of Information

In this section, I identify the conditions under which it exists an equilibrium in which all signals are credibly revealed. More precisely, I determine the conditions under which there exists a Perfect Bayesian Equilibrium Equilibrium where

$$m_{ij}(x_i) = x_i \text{ for all } i \in N,\ j \in N_i(g),\ x_i \in \mathbb{R}$$

Notice that, should all information be revealed, there exists a linear equilibrium in the last stage[11], where actions will satisfy.

$$a_i = b_{ii}x_i + \sum_{j \in N_i(g)} b_{ij}m_j = \sum_{j \in N_i^*(g)} b_{ij}x_j$$

for some weights where $b_{ij} \geq 0$ and $\sum_j b_{ij} \leq 1$. In general, the weight that i puts on signal x_j will depend on the total precision of the report of agent j - that is, the accuracy of both its signal and its message -, the total amount of information i has access to and the weight that others put on that signal[12]. Suppose $ij \in g$ and consider the incentives of agent i to truthfully reveal his type to j whenever everybody else does so. Using the envelope theorem, I write his indirect utility in the last stage as

$$\begin{aligned} -E\left[V_i\right] &= E\left[\Big(\sum_{j \in N_i^*(g)} b_{ij}x_j - \theta\Big)^2\right] + \\ &\qquad \frac{1}{n-1}\sum_{k \neq i} E\left[\Big(\sum_{h \in N_i^*(g)} b_{ih}x_h - \sum_{l \in N_j(g)\backslash\{i\}} b_{jl}x_l - b_{ji}m_i\Big)^2\right] \end{aligned}$$

[11] See Lemma A.2 a proof of existence of linear equilibrium. Notice that if communication is truthful our model reduces to a standard beauty-contests with agents receiving a number of signals equal to their degree and with endogenously determined precision. See Myatt and Wallace (2011) for such a model.

[12] See Lemma 1 in Appendix .2

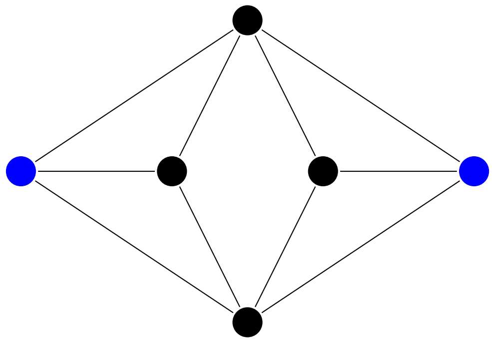


FIGURE 1. A Balanced network

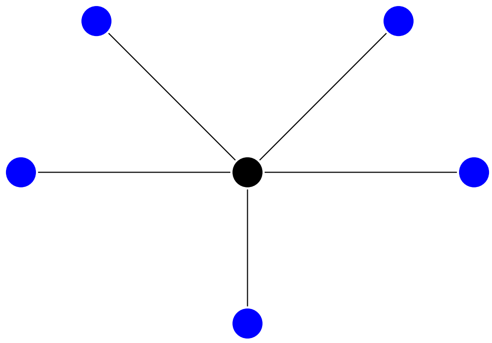


FIGURE 2. A Star network

First order condition for truthful revelation is then

$$
\begin{aligned}
(2.2)\qquad b_{ji}x_i &= E\left[\sum_{h\in N_i(g)} b_{ih}x_h - \sum_{l\in N_j(g)\backslash\{i\}} b_{jl}x_l \mid x_i\right] \\
&= b_{ii}x_i + \left[\sum_{h\in N_i(g)} b_{ih} - \sum_{l\in N_j^*(g)\backslash\{i\}} b_{jl}\right] E\left[x_k \mid x_i\right]
\end{aligned}
$$

DEFINITION 2.1. A networked information structure $\{g, \tau\}$ is **balanced** if for every i, and for every $j \in N_i(g)$, such that $\tau_i + \tau_j > 0$

$$
(2.3)\qquad \sum_{l\in N_i(g)} \tau_l = \sum_{k\in N_j(g)} \tau_k
$$

Thus, a networked information structure is balanced if for every two agents who communicate, the amount of information they expect to receive before taking their actions, and so their hierarchies of beliefs are ex-ante aligned. Regular networks where every agent acquires the same amount of information or star networks in which only the center acquires information satisfy this conditionv (see Figure 2). Another such network is depicted in Figure 1, where black nodes are informed and blue nodes are uninformed. The following Proposition shows that if a networked information structure fails to satisfy it, no equilibrium involves truthful information transmission.

PROPOSITION 2.2. *There exists an equilibrium with truthful revelation at every relevant link if and only if either $\tau_\sigma = 0$ or the networked information structure is balanced.*

PROOF. The proof is a straightforward application of the results in the literature. If $\tau_\sigma = 0$, $E[x_k \mid x_i] = x_i$ and $\sum_{j \in N_i(g)} b_{ij} = 1$ for all $i \in N$. Hence condition (2.2) holds. Otherwise it is needed that

$$(b_{ji} - b_{ii})x_i = \left[\underset{h \in N_i(g)_{ih}}{b} - \sum_{l \in N_j^*(g) \setminus \{i\}} b_{jl} \right] E[x_k \mid x_i]$$

Clearly, this holds if the network is balanced, because $b_{ji} = b_{ii}$ and $\sum_{l \in N_j(g)} b_{jl} = \sum_{k \in N_i(g)} b_{ik}$. Now, assume that the network is not balanced and suppose that the exists a pair $ij \in g$, $|N_i(g)| > |N_j(g)|$, I shall show that if the network is balanced j cannot make any information acquisition effort and that $\sum_{l \in N_{i-j}} \tau_l = 0$. In particular, and without loss of generality, assume that $N_i(g) = N_j(g) \cup \{k\}$. Notice that since the network is balanced it must be the case that $\tau_k = 0$. However, since $\sum_{l \in N_i(g)} \tau_l = \sum_{l \in N_i(g)} \tau_l$, if $\tau_i > 0$, there must exist and $k \notin N_i(g)$, there must exist another $k' \in N_k(g) \cap [N \setminus N_j(g)]$ such that $\tau_{k'} = \tau_i$. However, again $\sum_{l \in N_{k'}(g)} \tau_l = \sum_{l \in N_i(g)} \tau_l$. Hence, either $\tau_j = 0$ or $|N_{k'}(g) \cap [N \setminus N_j(g)]| > 1$. Finally notice that the total amount of agents is assumed to be finite, so that there must exist some k^t such that $N_{k^t} \subset N_{k^{t-1}}$ and hence, a contradiction with the network being balanced. Therefore either $\tau_i = 0$ or $\tau_j = 0$, so that $b_{ji} = b_{ii}$.

To see that it never holds if the network is not balanced notice that b_{ji} the only source of discrepancy between players is the amount of information received. In particular, i and j agree about the degree of agent j so that there is no bias generated in asymmetric networks per se. However, if i holds more information than j

$$\frac{\tau_i}{\sum_{k \in N_j(g)} \tau_k} > \frac{\tau_i}{\sum_{k \in N_i(g)} \tau_k}$$

and so $b_{ii} \neq b_{ji}$. But then

$$\sum_{h \in N_i(g)} b_{ih} - \sum_{l \in N_j^*(g) \setminus \{i\}} b_{jl} \neq 0$$

so that truthful revelation will not be part of any equilibrium. □

The intuition for the result is simple. Truthful revelation requires that hierarchies of beliefs are ex-ante aligned. This holds if the prior does not convey any information or if information is symmetric (total precision of the signals received by every agent is the same.) The reason is that an informative prior creates a wedge between the expectation of the underlying state conditional on a given signal and the signal. Hence, second order beliefs - the belief of i about the belief of j about θ- conditional on i's signal will differ with the current belief of i. This generates an incentive to i to lie her information and align those beliefs.

Notice that in our model cheap-talk equilibria does not rely on an *exogenous preference bias*. Ex-post, all agents would be better off if they had communicated perfectly their signals. However, in the interim, if the networked communication structure is not balanced, they have an incentive to misreport their information. In particular, well-connected agents have an incentive to make *conservative* reports about the state of the world, while badly connected ones have incentives to make *aggressive* reports. In equilibrium, these biases are understood by the receiver and they result in a reduction of the amount of information conveyed. Thus, differences in ex-post information, introduces vagueness in communication and reduces welfare.

Every equilibrium in the information transmission game is characterized by a partition of the set of signals, where a given report m is to be understood simply as $x \in [m_k, m_{k+1}]$. However, given the Gaussian structure of our uncertainty the equilibrium of this game is not tractable. In general, networks in which agents with different information communicate, are unable to share all their private information. This generates a twofold effect on the marginal value of information. First, it encourages information acquisition since agents have access to less information. Second, it discourages information acquisition since they lose ability to affect other players' behavior and thus reduce the coordination value of information. In this sense, information is less valuable if it is difficult to transmit.

In other words, if a network can attain an efficient communication equilibrium, information does not *endogenously* depreciate in its spread through the network. More generally, the degree of depreciation (or substitution among signals) will depend on the amount of information acquired by each agent and the degree distribution of the network[13].

Let $T = \{i \in N : \tau_i > 0\}$ be the set of agents who acquire information. Define $g^T = \{ij \in g : ij \cap T \neq \emptyset\}$ as the connected component of $g \cap T$ in which all the links which did not allow any information flow ($\tau_i + \tau_j = 0$) are deleted.

[13]See Section 4 for a more detailed analysis on the amount of information acquired.

COROLLARY 2.3. *Assume that* $g_k^T \subset g$ *is a balanced component of* $g \cap T$. *Then, there* $\exists J \subset N_k^T$ *and* $\gamma > 0$ *such that* $\tau_i = 0$ *if* $i \notin J$ *and for all* $i, j \in J$

$$\left|N_i^*(g^T) \bigcap J\right| = \left|N_j^*(g^T) \bigcap J\right| = \gamma$$

and if $\gamma \neq |J|$ *then for all* $i \in J$, $\left|N_i(g^T)\right| = \psi$, *for some* $\psi \geq \gamma$

Notice that the corollary implies that each component of a balanced networked organization can be partitioned in two subsets of agents. One of these subsets (J) contains all those who acquire information. Every player must be linked to a given number of agents within this set (γ), potentially including himself. This class of networks include the regular network, the core-periphery and mixtures of both. Figure 1 represents one of the possible configurations satisfying 2.3. Informed agents are depicted with a black node, while uninformed agents are depicted in blue.

This Corollary does also restrict the set of information acquisition patterns for other type of networks. For instance, in a line this condition requires that it can be partitioned in $2 + (N - 5)/3$ pairs of players who acquire information linked by agents who do not acquire information. Moreover, if every agent acquires information, there is no agent who can communicate truthfully to all of his neighbors. Indeed, we have the following result[14].

PROPOSITION 2.4. *Let* g^T *be a line with more than three individuals and assume that* $\tau_i > 0$ *for all* i. *For all* $i \in g^T$, *there exists* $j \in N_i(g^T)$, *such that* i *cannot communicate truthfully with* j

PROOF. In Appendix A.2 □

A line is depicted in Figure 3. Terminal nodes (depicted in blue) are disadvantaged in terms of the information they expect to receive. Therefore, when communicating they tend to be very aggressive and thus, equilibrium communication is vague. Because of this, their neighbors now are disadvantaged when communicating with others and are very aggressive with their neighbors. Following this reasoning one can show that vagueness in communication between two agents (under some conditions on the network structure) implies that vagueness will spread through every node of the component. This gives a new rationale for dense networks in organizational design, since adding links may not only increase the amount of information

[14]This result could be extended by allowing for more general networks although some conditions are necessary. In particular take g to be the complete network and delete the link $\{ij\}$, it is easy to see that among the rest of agents communication does not change. On the other hand, it is clear that symmetric trees could be included in this Proposition (a symmetric tree is such that all agents have the same number of links except the terminal nodes, who only have one)

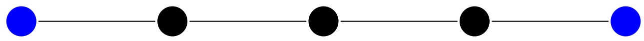


FIGURE 3. A Line

transmitted through new connections but also soften the incentive constrains in the communication with previously existing links.

2.4. Information Acquisition and Communication

So far we have seen that regular and core-periphery networks can support efficient communication equilibria for a given pattern of information acquisition. However, we would like to know whether indeed the pattern of information acquisition would lead to such networked information structures. That is, starting with a given network g, is there an equilibrium in the information acquisition stage-game such that the resulting $\{g, \tau\}$ allows for truthful communication?

PROPOSITION 2.5. *Assume that the networked information system g^T is connected and that communication is truthful in the continuation game. Then, τ_i and τ_j are strategic complements if and only if $j \notin N_i(g)$. Alternatively τ_i and τ_j are strategic substitutes if and only if $j \in N_i(g)$.*

PROOF. In Appendix A.2 □

This result highlights the importance of studying the interaction between communication and information acquisition, even in the case of truthful information transmission. The literature on information acquisition in Beauty Contests have shown that if there is need for coordination, information acquisition satisfies strategic complements. On the other hand, if there is communication, information becomes a public good and information acquisition must satisfy strategic substitutes. Finally notice that a similar result holds in Calvó-Armengol et al. (2011) but in their framework each agent is concerned with the realization of a local random variable and the information of others is only relevant to predict their action.

The literature on communication networks has also highlighted the role of centrality (see Calvó-Armengol et al. (2011)) in shaping up the incentives to acquire information. Does this monotonic relation extend to this framework? A negative answer is provided below

EXAMPLE 2.6. Assume that g is a line with $3n+5$ players with n integer, and let $c(\tau) = k\tau$. There exists an equilibrium g^T with truthful information transmission and non-monotone pattern of information acquisition

The idea for this counterexample is depicted in Figure 4. An equilibrium exists with truthful revelation of information, where only those nodes depicted in black acquire information. To see that this

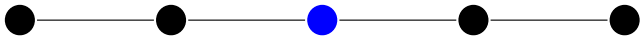


FIGURE 4. A line with truthful revelation

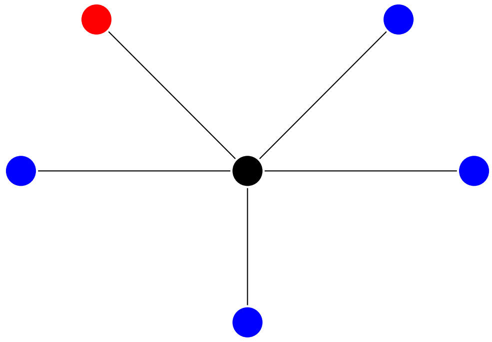


FIGURE 5. A Star network

is an equilibrium under the prescribed technology, notice that blue nodes have lower value of information than black nodes (since they will trigger communication losses from any marginal increase in information acquisition) and face the same marginal cost. Thus, they optimally restrain from acquiring information.

However, there are certain conditions under which the result obtains. Let $V_i(\tau)$ be the expected utility of an agent located in position i if the vector of information acquisition patterns is τ. Thus if g satisfies Corollary 2.3, a necessary condition for a τ^* equilibrium profile to be such that $\{g, \tau^*\}$ is balanced is that

(2.4)
$$\frac{\partial}{\partial \tau_j} V_j(\tau^*) - c'(\tau_j) \geq 0 \geq \frac{\partial}{\partial \tau_i} V_i(\tau^*) - c'(0), \text{ where } j \in T^*, i \in N \setminus T^*$$

COROLLARY 2.7. *Assume that $c(\tau)$ is linear. If g can be partitioned so that $g^T = \cup g_k^T$, where for each $k = 1, 2, ...K$, with $0 < K \leq N$, and g_K^T satisfying the conditions in Corollary 2.3. Then, there exists an equilibrium truthful information transmission.*

The intuition behind the Proof of this Corollary is depicted in Figure 5. If the player located in the red node increases his information acquisition effort, and the cost is linear he would obtain an increase in his payoff not larger than the increase that the central node would obtain. This is because more precise signals for the red node worsens communication at every other link and diminishes total coordination. Thus, specialization obtains where only the central node acquires information.

2.5. Welfare and Information Acquisition

There are two important questions for which we do not have a definite answer. First of all, if we do not have full information transmission, can we still say something about the pattern of information acquisition? In particular, it is important to know whether more central agents will still acquire more information despite the fact that they cannot communicate it perfectly. The following Proposition shows that this also fails to be true.

PROPOSITION 2.8. *Let a component $g' \subset g^T$ be a connected line, and assume that $\tau_i > 0$ for all $i \in g'$. There exists some $\hat{n} > 3$, such that if $n' > \hat{n}$, extreme players acquire more information than their neighbors.*

This follows from the fact that if the residual uncertainty of the extreme player is always larger than or equal to that of the extreme one, and if the network is sufficiently large, this implies higher incentives for information acquisition.

In general, different communication structures would lead to different patterns of information acquisition. In many studies information acquisition is treated as any other public good (see, for instance, Galeotti and Goyal (2010)) so that, the signals one agent collects and the signals others collect are perfect substitutes. However, in our model, the degree of substitutability is *endogenous*. In particular, if communication is truthful, those signals received by the neighbors of i are perfect substitutes from the signals coming from j[15]. This implies that the effort of i decreases in the effort of his neighbors. On the other hand, if the communication equilibrium were characterized by full babbling, information acquisition efforts become strategic complements (see Hellwig and Veldkamp (2009)) Our strong conjecture is that there is a monotonicity in the degree of substitution between information acquisition strategies of different players and the quality of their communication. This link highlights the importance of considering both problems simultaneously.

Another important issue is that of welfare. That is, which networked information structures yield better outcomes in terms of coordination and adaptation? Clearly, among minimally connected networks, either the star or the circle can be optimal. The star provides better coordination but only one agent acquires information. In a circle, every agent aggregates two signals which may offset the coordination losses from decentralized information. This depends on the cost of information acquisition and the environmental uncertainty.

A somewhat more important question is whether adding links does always lead to a reduction in the total variance resulting in

[15]Note that in this model, Gaussian signals imply that all that matters is total precision. See Proposition 3.

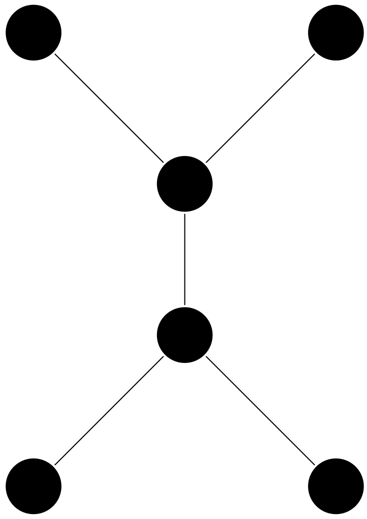


FIGURE 6. Welfare.

decision-making [16]. The following example shows that this is not the case

EXAMPLE 2.9. Suppose that g^T is formed by two star components with sufficiently many agents and the most revealing equilibrium is being played. Adding the link between both hubs is welfare detrimental.

An example of such network is depicted in Figure 6[17]. In the most revealing equilibrium of the game generated by the original network, two agents acquire information and they communicate it truthfully to their neighbors, each of which receives one signal and suffer small coordination losses (they fail to coordinate with the other component.) If both hubs link to each other, now in any equilibrium welfare (as measured by total variance) is lower. First, assume that the pattern of information acquisition does not change. Now, coordination only improves between the hubs and worsens between every other linked agents. Further, adaptation also decreases at every other player. Thus, if there are sufficiently many agents, the sum of ex-ante payoffs decrease. Now it may also be that either 2 or 6 (or both) cease to acquire information. In this case the comparison is even simpler. Each of the agents linked to the hub who ceases to acquire information would increase his information acquired and there would be wasteful duplication of effort, while each of the spokes would still obtain one signal.

[16]This criterion is used in the team-theoretic literature starting from **?**

[17]more peripheral agents may be added if needed

2.6. The Role of Hierarchies

Most of the previous literature on communication in organizations agreed that hierarchies are an efficient way to transmit and process information. Radner (1993) shows that a hierarchical structure (a tree in the jargon of graph theory) is the most efficient structure for an organization that tries to process and summarize a large amount of disseminated information. Bolton and Dewatripont (1994)extended this logic to environments with an infinite stream of signals that have to be processed with minimal delay. Geanakoplos and Milgrom (1988) showed that, under bounded rationality of managers, hierarchical organizations are the most efficient way to use a group to overcome the limitations of its members. Garicano (2000) shows that a hierarchy is the natural organization for a firm that must solve problems in order to produce if workers cannot identify those problems that they cannot solve.

A common feature of all these models, however, is that they assume that the members of those organizations are not strategic. In particular, they acquire the information they are told to acquire, they transmit it truthfully and they take the action that the organizations wants them to take, conditional on the information available. In our model, however, agents are rational, strategic players who try to maximize their payoffs in a coordination game under uncertainty. Uncertainty creates a wedge in the way agents with different locations update their beliefs and, therefore, incentives to misreport their information. In hierarchies (or trees) agents at the top are bound to receive more information than agents at the bottom, and thus, information transmission fails to be efficient.

In the real world this problem is solved in the following way. Information is acquired by lower-ranked agents who communicate it upwards to managers. These managers aggregate information and pass it back to the periphery in the form of ”recommendations” or ”commands”. Therefore, although hierarchies are efficient in terms of information handling they require some source of ”power relation” among agents in order to conveniently achieve the organizational goal. Our model lacks this power relation since every agent is entitled to ”decide” and information does not move backwards once it has been aggregated.

2.7. Repeated Communication

One of the main driving forces of the results presented above is the use of interim beliefs and one round of communication. That is, since agents only communicate once, they rely heavily on the beliefs they hold after observing their signal, when making their reports.

This is the source of their *intrinsic bias*. I shall explore now this assumption by constructing a dynamic environment in which the game presented above is (potentially) infinitely repeated. I construct the game following the ideas in Acemoglu et al. (2010)).

The game is as follows. At time zero, every agent makes some investment in information. Then, both θ and the signals are drawn from the appropriate distributions. At time $t = 1$, agents report through the network $g^1 = g^T$ some messages conditional on their signals and their positions in the network $m_{ij}^1(x_i) \in \mathbb{R} \cup \{\emptyset\}$ [18]. Let m^{i1} be the profile of reports received by agent i in period 1. Then, agents take actions $a_{i,1} \in \mathbb{R} \cup \emptyset$ where $a_i = \emptyset$ is defined as inaction. After that, actions are realized and agents who took an action leave the game[19]. Agents who decided to stay inactive keep move into the next period by losing $\delta > 0$. At time $t = 2$, $g^2 = g^1 \cap \{i \in N : a_{i,1} = \emptyset\}$ and again chose a report $m_{ij}(x_i; m^{i1}) \in \mathbb{R} \cup \{\emptyset\}$ where $ij \in g^2$ and an action $a_{i,2} \in \mathbb{R} \cup \emptyset$. Whenever at the end of time t, the set $\{i \in N : a_{i,t} = \emptyset\} = \emptyset$ the game ends and every agent receives his payoff according to the original payoff net of the corresponding loss for delay δt_i, where $a_{i,t_i} \neq \emptyset$.

PROPOSITION 2.10. *Suppose that $\{\tau, g^1\}$ is balanced. Then there is an equilibrium of the repeated game in which every agent reports truthfully and makes an action in the first period.*

This implies that the positive results in the previous section survive into this extended game. However, do the negative ones survive? A qualified answer would be yes. In particular, there are some networked information structures at which information cannot be transmitted at *any* round.

PROPOSITION 2.11. *Suppose that $\{g^1, \tau\}$ is a line, and assume that $\tau_i > 0$ for all i. Then for every $t = 1, 2, ..\bar{t}$, there exists an equilibrium in which all agents leave at period t. Further, no equilibrium involves perfect communication.*

PROOF. In Appendix A.2 □

The intuition is simple. Suppose that a given agent (a) has only one neighbor (n). Suppose further that his neighbor has at least one additional informed neighbor. n takes the action (and leaves) one whenever he has acquired enough information, and therefore does not transmit the last piece of information to a. Therefore, at every stage before exit, n has an incentive to misreport his information to

[18] We allow for explicit witholding of information.

[19] This is the main restriction of the framework since it will not be optimal for them to leave (for sufficiently small δ) and clearly their information is still valuable for others. However, since their actions are now taken, there is no hope that their reports are truthful. Indeed, it is easy to construct examples in which overall welfare is lower if all agents stay until everyone took their actions.

a. Thus, in this extension of the game, as long as players leave the network after taking their actions with positive probability the qualitative features of the static equilibrium remains, even if they hinge on agents using interim beliefs.

In other words, whenever communication takes place between agents who ”expect to learn more on the future”, our results are likely to hold. However, in most studies, the assumption is that players are either informed ex-ante or ex-post but they never get ”some information” in the interim. Repeated communication games, for instance, assume that players have acquired all the relevant information at stage zero. We argue that this assumption has deep implications in the results, and it is not clear why this possibility should be ruled out.

An exception in this literature is Acemoglu et al. (2010) They allow for strategic communication of signals using interim beliefs. However, in their model the strategic interaction is different, since a given player does not care about the action taken by others. In particular, an agent would only want to lie if that induces other agents remain in the network and keep communicating with him. To manage so, she is willing to misreport her true signal in order to ”confuse” her peer and make him stay. This is not possible in our model since the residual variance does not depend on the ”content” of the reports.

PROPOSITION 2.12. *Suppose $\{g^1, \tau\}$ is not balanced, and assume that $\tau_i > 0$ for all i. If there exists $i \in N$ such that $N_i(g) \subset N_j(g)$, $N_i(g) \neq N_j(g)$, then either both agents hold the same information (and take the same action) or there is no equilibrium with perfect information transmission between them.*

PROOF. In Appendix A.2 □

This result shows that poorly informed individuals (in that they have access to a subset of the sources of their neighbors) will also have problems to communicate with those sources, independently of how many rounds of communications are allowed. This may seem counterintuitive, since an agent who talks to a poorly informed agent has a very good posterior belief over the belief of his neighbor. The problem is that this second order belief may be far away from the belief he holds! This yields a novel intuition that was not present in the static game. Namely, since agents are willing to wait only if waiting yields new and useful information, communicating with agents who have access to that information is not only more useful, but easier (in the sense that it reduces the vagueness in communication) than to those who have no new information.

Equilibrium behavior depends (discontinuously) in the discounting of players, as it is the case in many dynamic games. This is because the value of staying in the network depends on the number of neighbors who remain and communicate their signals. If the cost of continuing in the game is high enough, some (informed) players decide to leave early and communication does not spread. If the cost is low enough, however, the network structure is not relevant since information would travel in a frictionless manner and eventually coordination would be achieved. In this sense, it is the cost of time that gives a specific content to the network itself.

2.8. Conclusion

2.8.1. Summary. In this chapter I argue that modeling explicitly the information acquisition and transmission may be important to understand the functioning of many organizations and markets, even if there are no exogenous conflict of interest between the players. I show how the topology of the communication network and information acquisition technology affect the quality of information transmission within the organization. I fully characterized the set of networked information structures that support perfect communication as an equilibrium and the pattern of information acquisition they generate. These networks have the property that every two individuals who communicate are expected to hold the same amount of information by the time they make their decision.

I also show that whenever information revelation is not truthful, the pattern of information acquisition effort may change dramatically. For instance, in the line with sufficiently many players, no agent can communicate truthfully with all of her neighbors and the information acquired in equilibrium is not monotonic in centrality. This result is in sharp contrast with the previous literature, which highlighted the role of centrality in the intensity of the effort.

I have also extended the model to allow for more rounds of communication. I have shown that our results are qualitatively robust to such an extension, as long as players abandon the network as soon as they make their decisions. If an agent expects to receive more information in the future, she will use her interim beliefs when forecasting the actions that their communication partners are going to undertake. If she expects to received a different amount of information in the future (compared with her communication partners) she will be biased, and truthful communication will not be an equilibrium.

Finally, the analysis has provided insights on the public good nature of information in networked organizations. First, whenever communication is not truthful, information transmission exhibits decay or endogenous depreciation. This depreciation depends on the density of the network because more links reduce average distance between two randomly chosen nodes and also because it allows for better information transmission. Second, in most organizations, the amount of information acquired by one agent may be increasing or decreasing in the total amount of information acquired by others. In a setting where information flows perfectly, information would be non-rival and the effort in information acquisition would become strategic substitutes. However, in our setting agents need to ”be like their peers” in order to extract information from them. Moreover if information transmission is noisy, they need to ”know what others know” in order to predict their behavior.

2.8.2. Discussion and Future Research . The main limitation of this chapter is the inability to obtain close form characterization of the equilibrium under imperfect information. This precludes any analysis on the effect of imperfect communication on the equilibrium decisions or welfare comparison across different networks. These issues, albeit important, have already been studied in depth in rational-expectations models[20]. In this sense, this analysis complements this strand literature, by introducing endogenous information acquisition and strategic communication. These papers find that more centralized networks lead to excess volatility in prices and that denser networks lead to excess correlation across agents. Interestingly, they have devoted most of their attention to those network topologies that our model predict to be more efficient in information transmission.

As indicated earlier future research may shed some light on the information transmitted between players who cannot commit to truthfully reveal their information. One approach would be to allow only for all-or-nothing communication, i.e. any two links can sustain either truthful revelation or no communication at all. Another approach would be to modify the structure of the model and introduce, for instance, a simpler, discrete state space, where the information revelation strategies can be pinned down. These extensions would allow for a much clearer rank of networked information structures.

[20]See Xia (2010) and Ozsoylev and Walden (2011)

CHAPTER 3

Information Technology, Decision Making and Incentives

> 'Everybody's marriage is falling apart except ours. You see the problem is communication.... too much communication.'. Homer J. Simpson

3.1. Introduction

One of the main objectives of the very existence of organizations is the use of sparse information to achieve a common objective (Arrow (1974)). Modern organizational theory has emphasized the role of private incentives in the acquisition, transmission and use of information. According to this literature, private incentives may not be fully aligned with the common goal, inducing inefficiencies in decision-making. To solve this problem, a number of organizational design tools like contingent compensation and promotions or delegation have been proposed. In this chapter, I argue that another way in which some of these problems may be addressed is the communication and information system itself. The way an organization uses and transmits information would shape and (potentially) align the incentives and behavior of its members, and in this sense contribute to the creation of value for the firm.

As a matter of example, consider the problem faced by a global company selling some product in different countries. Each national unit would like the product to be designed for their particular demand characteristics, but making a product that is fundamentally different for every country may preclude the organization to obtain cost economies. Adaptation to local conditions entails higher sales and profits for every unit but limits coordination and may harm the overall performance of the company. To solve this problem, decisions made by the central unit would not fully incorporate local characteristics, which on its turn creates an incentive for each particular unit to overstate their need of adaptation. Whereas some recent studies[1] have analyzed this problem using a cheap talk game[2] we allow each

[1]See, for instance, Alonso et al. (2008) and Rantakari (2008)

[2]That is, a game in which an agent makes an unverifiable report concering the state to an informed party that makes a report-contingent decision about some variable that affects the payoff of both players.

local party to perform an investment, tailored to the payoff-relevant state (e.g. demand characteristics). In our example, these investments may be thought of as the advertising and marketing effort of the national units to sell a product that is not very well suited to their characteristics. The important aspect is that different demand characteristics require different levels of investment, so their observation allows the central unit to infer the true state of the world. The main question that we can address in such a framework is which is the optimal precision with which the central unit would like to learn the investment decisions by each unit. That is, which factors underlie the adoption of more efficient information technologies (IT).

There is more than anecdotal evidence that this is an important problem in many organizations. Management scholars had long recognized the problem that headquarters face whenever they have to rely on information from their business units. Chandler's study on the management practice of some of the biggest US firms in the first half of the 20th Century devotes significant attention to these problems (Chandler (1969)). In particular, both Standard Oil and General Motors struggled to obtain both accurate information and appropriate coordination from their Business Units. As Chandler points out for the case of Standard Oil[3]

> "(...) But any allocation of funds or other resources was based on the information provided by the affiliates themselves. The parent made little effort to check on the way they employed their resources"

To model this environment, we use a three-stage game played between the business units and the headquarters (HQ). In the first stage, the central unit decides the information technology of the firm. That is, she chooses the quality of communication and the precision of the signals that will be transmitted inside the organization. In terms of the previous example, each agent may submit a report to the central unit, specifying their local characteristics and the investments they will undertake to improve their payoffs. The precision with which the central unit learns this may depend on the information technology available and the resources allocated in this process. We view this as a long-term decision which give some commitment power to the organization, and explore the role of this power in softening the strategic incentives to invest. In the second stage, each party learns their type and (subsequently) undertakes a costly investment. Both the state of the world and the action determine the preference relation over the possible decisions that the principal may choose. Finally, in the last stage, the principal observes each report through the chosen partition and makes a decision for each party.

[3]See Subsection 3.1.2 for a more detailed case

3.1.1. Related Literature. The present chapter is related to a number of strands in the literature. Inspired by the seminal work of Marschak and Radner (1972), recent works by Alonso et al. (2008) and Rantakari (2008) analyze the conflict existing in multi-divisional organizations that must adapt to local conditions. Agents communicate via cheap talk messages between themselves and with the central unit. Surprisingly, the incentive conflict is better solved by relying on horizontal rather than vertical communication. Thus, they obtain that decentralized decision making (which in their framework implies giving decision rights to the local parties) may be superior to centralization even if coordination is very important. As pointed above, the main contribution of our chapter stems from the introduction of (observable) investments that may help soften the incentive conflict within the organization, and the analysis of the way decision-making and investments will be distorted because of the agency problem these organizations face.

Many papers have recently studied the optimality of various decision-making schemes in delegation problems. Szalay (2005) shows that the principal may distort ex-post decisions to induce the agent to acquire relevant information. Ambrus and Egorov (2009) show that inefficient (money-burning) activities may be useful as a way to provide incentives for more efficient decisions. Ivanov (2010) studies a standard delegation problem in which the principal may control the precision of the information received by the agent. He shows that the principal will be better off by restricting the precision as a way to align incentives. Both studies interpret their results as a rationale for more centralized decision-making processes[4].

Other related papers that study the problem of decision-making and incentives include Athey and Roberts (2001) and the recent paper by Dessein et al. (2010). In their framework, agents are delegated decisions concerning overall profits but must be rewarded using only a partial measure (because of a multitasking problem.) The main difference is that the effort is related to a different task, and hence it generates the incentive misalingment.

Finally, I would also like to point that this chapter relates to a (small) strand of the organizational design literature in which the object of study is the information and communication technology. Hansen (2010) studies a career concern model in which the decision of how much feedback to give to workers is endogenous. He points out a similar trade-off between information and incentives in a signaling game, but he does not consider the effect on decision-making,

[4]This literature interprets lower response of decisions to the environment to higher centralization. In our model the optimal degree of centralization is determined by the environment, but the variation of decisions will depend on the endogenous information structure.

where information is useful. Lee and den Steen (2010) have a dynamic model in which the firm chooses how much information about technological process to store to trade-off innovation and efficiency.

3.1.2. The Case of Novo Nordisk . Novo Nordisk [5] is a Danish Pharmaceutical Corporation whose main business is the production and commercialization of insulin. It is the market share leader in most countries and the second company world-wide. Insulin is a very standardized product, and the technology used presents huge scale economies - one factory could account for a third of the world production. Most of the innovations are related to the ease of use of the product and the reduction of secondary effects, to facilitate commercialization and compliance to changing regulations.

Novo Nordisk produced all their insulin in Denmark but packed it and prepared it for human use in different countries, according to their specific needs. Regulations and distribution channels are indeed the most important piece of local information that this firm needed to succeed. However, local units empowerment was weak and although communication was fluid and the HQ received a lot of information about their performance, they rarely were held accountable of their actions. In particular, all that was required was that the local unit complied with the main values of the firm.

This seemed to work fine in most of the world but, apparently, it was not the case in the US. Their main rival, Eli Lilly, was clearly ahead in that market, and the HQ claimed it was mainly the local unit fault. On their turn, the US branch claimed that they were not given enough freedom and that legal conditions in the US were very different. Further, they were claiming that F.D.A. - the American Agency for Food and Drug certification - were increasing there controls and restrictions after a war between generic and non-generic producers in the US.

In this environment of conflict, Novo Nordisk tried to gain ground on their rivals by introducing a new product that enabled patients of type II diabetes - the less severe one - to control it in a much more efficient and comfortable way. However, the F.D.A. did not approved this new product since its production did not meet the new standards they impose. Novo Nordisk decided to stop the production worldwide and defer their clients in the US to their main rival, with considerable losses. HQ understood that new managerial practices were called upon.

Although they recognized this as a coordination problem, they did not call for higher centralization. On the contrary, they increased empowerment of the international divisions. Further, they eliminated

[5] This Subsection is based on a Podolny et al. (2000)

many of their information feedback programs, and decided to focus on simple statistical measures to appraise performance. These changes seemed to have positive impact on performance, since they managed to recover their position by 1997.

We would like to highlight that this case fits well into our model and allows us to anticipate some of the main results. In particular, since the US division could have pushed the F.D.A. towards a different policy, it is clear that local information was interrelated with the performance of the different division, which is our main departure from previous literature. Second, the solution proposed to achieve better coordination implied less information transmitted but better incentives to provide the appropriate level of effort. Third, increasing empowerment did not lead to higher statistical control but to an increase in the responsibility of each unit.

3.2. The Model

As we pointed out above, we will consider organizations composed by a central unit (p) and N agents (i). Each agent observes a realization of a random variable $\epsilon_i \sim F(\epsilon)$ a symmetric distribution with full support on $[-\bar{\epsilon}, \bar{\epsilon}]$. Let f be the associated density. We shall use $\epsilon = (\epsilon_1, \epsilon_2, ..., \epsilon_N) = (\epsilon_i, \epsilon_{-i})$. The agent may perform an effort $a_i \in A$ some compact and convex subset of $\mathbb{R}$ containing $[-\bar{\epsilon}, \bar{\epsilon}]$, with cost $c : A \to \mathbb{R}_+$ symmetric and (weakly) convex with $c(0) = 0$. Let $a \in A^N$.

The principal chooses the informational system for the firm in the first period. Formally, the central unit chooses a correspondence $m : A \to 2^A$ with a typical element $m(a_i) = A_k$ for all $a_i \in A_k \subset A$. We assume that whenever A_k are non degenerate, A_k are Borel sets. Notice that information is truthful but may be imprecise. Given the symmetry of the model, I will impose that information structures are also symmetric around $a_i = 0$. That is, if we order sets by their infimum, we have that $a_i \in A_{k-l}$ and $0 \in A_k$, then $-a_i \in A_{k+l}$. Once the information structure is chosen, all agents learn their types and choose an action from the set of actions. Finally the principal observes the actions performed according to the partition chosen and the report and makes sequentially optimal decisions. In the last stage, payoffs are realized.

Both the shock and the action determine the preferred point of the agent $\theta_i = \epsilon_i + a_i$ for the decision $d_i \in Y \subseteq [-\bar{\epsilon}, \bar{\epsilon}]$ chosen by the principal. Then we can write the agent's payoff as

$$U(\epsilon_i, a_i, d_i) = -(\theta_i - d_i)^2 - c(a_i) \tag{3.1}$$

Notice that we abstract from monetary transfers, since participation constrains are not relevant for our framework, and contingent transfers are not allowed[6]. We assume that the principal cares about efficiency of the whole organization. In particular, he would like to maximize the sum of profits in each division but must bear a cost from adapting decisions to each unit, so we write its payoff as

$$V(d) = -\sum_{i \in N}(\theta_i - d_i)^2 - \gamma \sum_{i \in N}(d_i - \bar{d})^2 - \sum_{i \in N} c(a_i) \tag{3.2}$$

where $\gamma > 0$ measures the relative importance of coordination and

$$\bar{d} = \frac{1}{N}\sum_{i \in N} d_i$$

is the average decision. In this model there is no exogenous bias in the preferences of the agents with respect to the central unit. The incentive to misreport stems from the fact that the principal will respond to private information less than the agent would like to.

We would like to highlight that although its precise formulation is somewhat new, this problem belongs to the class of decentralization problems presented in section 2 of Holmstrom's seminal contribution (Holmstrom (1980)). We do not consider delegation since in our framework each individual decision will depend of the whole set of signals - not only on the signal received by that individual. Hence, we focus on communication procedures where the central party retains the formal rights to decide. In any case, we try to capture the spirit of delegation as an institution that requires low commitment. Whereas delegating principals commit not to overrule decisions, our principal commits to an informational system that operates in the firm and it is not expected to be changed in the near future. Notice that, as in the example above, the principal would like to commit to a different decision rule in the last stage, due to the interaction of private information and a free-rider motive that comes from the common decision. Allowing her to choose the precision of the informational structure somewhat softens that need but does not substitute it completely.

We now summarize the game generated by an information structure $m(\cdot)$, Γ_m. Observing the information structure and his realization, each agent picks one action $a_i \in A$ and the principal observes

[6]There are many reasons for this. First, the interest of this chapter is to study the role of information in decision-making and incentives. Including compensation would add more effects to this comparison. Second, linear-quadratic utility functions are not well-suited since it is the only single-peaked utility function where first best is attainable. Third, optimal compensation would be based on soft information, which would lead to verifiability problems.

$m = (m(a_i))_{i\in N}$, the profile of contingent reports from each unit. Finally the principal chooses some $d\in Y^N$. We concentrate on Symmetric Perfect Bayesian Equilibrium of the game starting after a choice $m(\cdot)$.

DEFINITION 3.1. A Symmetric Perfect Bayesian Equilibrium of Γ_m is a strategy profile for each agent $a_i^*(\cdot)$, a decision rule for the principal $d(m(a))$ and a set of posterior beliefs $\mu(\cdot \mid m)$ such that

$$\text{for all } \epsilon,\ a_i^*)(\epsilon) \in \arg\max E\left[U\left(\epsilon, a_i', d(m(a_i', a_{-i}^*))\right)\right] \tag{3.3}$$

$$\text{for all } m,\ d(m) \in \arg\max E_\mu \left[V(d'.a^*)\ |m\right] \tag{3.4}$$

$$\mu(\epsilon \mid A_k) = \frac{f(\epsilon)}{\int_{A_k} a_i^{*-1}(\ a)dF(\epsilon)},\ \text{for all } A_k \subset A \tag{3.5}$$

Solving it backwards, notice that in the last stage, the principal maximizes

$$V(d, a^*) = -E_\theta\left[\sum_{i\in N}(\theta_i + a^* - d_i)^2 \mid m\right] - \gamma\sum_{i\in N}(d_i - \bar{d})^2 \tag{3.6}$$

where the expectation is taken with respect to the distribution of payoff-relevant variables conditional on the information available in the last stage.

From (3.6) we have that

$$d_i(m) = \left(1 - \frac{(N-1)\gamma}{N(1+\gamma)}\right)E\left[\theta_i \mid m\right] + \frac{(N-1)\gamma}{N(1+\gamma)}E\left[\sum_{j\in N\setminus i}\theta_j \mid m\right] \tag{3.7}$$

which implies that the expectation of the decision for agent i conditional on his own report is.

$$E_{\theta_{-i}}\left[d_i(m) \mid m_i\right] = \left(1 - \frac{(N-1)\gamma}{N(1+\gamma)}\right)E\left[\theta_i \mid m_i\right] = (1-\lambda)E\left[\theta_i \mid m_i\right] \tag{3.8}$$

Given this decision rule, the problem of each agent is to choose an investment level to minimize

$$U(\epsilon_i, a_i, d_i) = E_{\theta_{-i}}\left[(\theta_i - d_i)^2 \mid m_i\right] + c(a_i) \quad s.t. (3.8) \text{ and } m(\cdot) \tag{3.9}$$

3.2.1. First Best. In this subsection we derive the first best level of effort. Notice that, without private information, the principal still chooses the same decision for a given profile of investment levels. In particular

$$d^{FB} = (1 - \frac{(N-1)\gamma}{N(1+\gamma)})(\epsilon_i + a_i^{FB}(\epsilon_i)) = (1-\lambda)(\epsilon_i + a_i^{FB}(\epsilon_i))$$

In order to make welfare comparisons, it is going to be useful to assume that $c(a_i) = ca_i^2$ for some $c > 0$. then, $a_i^{FB}(\epsilon_i)$ satisfies the following equation

(3.10)
$$a^{FB} = -\frac{(N-1)\gamma}{c(N(1+\gamma)) + ((N-1)\gamma)}\epsilon_i = -\frac{\frac{(N-1)\gamma}{N(1+\gamma)}\epsilon_i}{c + \frac{(N-1)\gamma}{N(1+\gamma)}} = -\frac{\lambda\epsilon_i}{c+\lambda} = \mu^{FB}\epsilon_i$$

where it holds that $\frac{da}{d\epsilon_i} \in (-1, 0)$. Notice that the slope is higher (in absolute value) whenever λ is bigger and whenever the cost is lower. We shall interpret λ as the need for coordination. It increases with the number of units (N) and with the relative importance of coordination γ.

3.3. Perfect Precision

We shall start the analysis considering the possibility that an organization chooses to perfectly observe the investment decision of its divisions. That is, $m(a) = m(a')$ if and only if $a = a'$. In our model inefficiency arises from the combination of unobservable actions, adverse selection and uncertainty. In principle then, one would expect that increasing precision would allow the firm to reduce both the informational advantage of its divisions and the coordination losses from sparse information. However, more information increases the possibilities of deviation by agents, thus tightening the equilibrium constrains and further aggravating the problem of a lack of commitment.

As it is standard in communication games, there is multiplicity of equilibria. First, there exists a fully separating equilibrium in which every type tailors his action to his shock and there is no ex-post uncertainty in decision making (see Proposition 3.2). However, as it is standard in communication games an infinite number of partial-pooling equilibria exist. In these equilibria, there is a positive mass of types who choose the same level of effort, and the set of investments that belong to equilibrium strategy is not (necessarily) connected. However, we show that this equilibria are always Pareto-inferior to the equilibria generated by choosing a different information structure, and hence, they are not relevant for the design problem of the firm[7].

Hence we shall consider first fully separating equilibria where every possible type makes a different investment level and hence, there

[7]We follow the convention in organizational design, that players coordinate on equilibria that are not Pareto dominated. See, for instance, Hermalin (1998)

is no ex-post imperfect information[8]. In particular, when choosing the level of effort, each agent takes as given the equilibrium strategies of the other players that, together with the decision rule of the principal, the mapping from actions to decisions. If the full separating equilibrium is to exist, then this mapping can be characterized as

$$d(a_i) = E_{\theta_{-i}}\left[d_i(a)\right] = (1-\lambda)E\left[\theta(a_i)\right] = (1-\lambda)\left[a_i + a^{-1}(a_i)\right]$$

whenever the inverse effort function is well-defined. If we concentrate in the quadratic cost case we can find a solution to the individual agent's problem (3.9) of the form $a_i(\epsilon_i) = \mu\epsilon_i - a(0)$. Because of the symmetry, it is going to be true that $a(0) = 0$, and so, $[a_i + a^{-1}(a_i)] = \frac{\mu+1}{\mu}a = \nu a$. In particular, first order condition is

$$2(\theta_i - d(a_i))\left[1-(1-\lambda)\nu\right] = -2ca \tag{3.11}$$

and so we can solve implicitly for ν as one root of the following equation

$$\nu - 1 = -\frac{c + \left[1-(1-\lambda)\nu\right]^2}{.}$$

So that

$$(1-(1-\lambda)\nu) = \frac{-c}{(1-(1-\lambda)\nu)} + (1-(1-\lambda)\nu)$$

$$\begin{aligned} \lambda\nu &= \frac{-c}{1-(1-\lambda)\nu} \\ \nu &= \frac{-c}{\left[1-(1-\lambda)\nu\right]\lambda} \end{aligned}$$

Since $c > 0$, it follows that there is a root with $\nu < 0$, and hence $\mu \in (-1, 0)$. The other root is positive and we discard it, since it cannot be optimal[9]. Notice that

$$\nu = \frac{-c}{(1-(1-\lambda)\nu)\lambda} = \frac{\mu+1}{\mu} \tag{3.12}$$

using (3.10) we have that $\nu^{FB} = \frac{-c}{\lambda}$, so if both cross it must be that

$$\frac{-c}{1-(1-\lambda)\nu)\lambda} = \frac{-c}{\lambda}$$

taking $c > 0$, implies that $\lambda = (1 + \frac{c}{\lambda}(1-\lambda))\lambda$. But since $0 < \lambda < 1$, they cannot cross. Further, notice that for any $\nu < 0$, $(1-(1-\lambda)\nu)\lambda > 1 > \lambda$ and hence, we have that $\nu^{FB} < \nu$. From (3.12), follows that $\mu < \mu^{FB} < 0$. That is, in the equilibrium with full separation, there is over-provision of effort, for any possible need for coordination. It follows that, no matter how hard is the public good problem, the signaling motive

[8]Notice then that condition B in the definition of Equilibrium becomes degenerate. This implies that only conditions P_1 and P_2 matter so that this equilibrium is Nash.

[9]Recall footnote 9.

generates over-investment. This is because, whenever coordination becomes more important, incentives become more misaligned and information transmission more inefficient. However, whenever coordination becomes more important with respect to adaptation, information is less valuable, so that information technologies are less important. In particular, using the equilibrium strategies, we can check that the marginal effect of reports on the final decision and effort provision are negatively related.

$$d\nu\left[(1-2(1-\lambda)\nu)\lambda\right] = -d\lambda(\nu-\nu^2) \tag{3.13}$$

$$\frac{d\nu}{d\lambda} = \frac{(\nu-\nu^2)}{\left[(1-2(1-\lambda)\nu)\lambda\right]} < 0 \tag{3.14}$$

$$\Rightarrow \frac{d\mu}{d\lambda} = \frac{-\frac{d\nu}{d\lambda}}{(\nu-1)^2} > 0 \tag{3.15}$$

Now we state this result in a proposition and show that indeed these strategies are an equilibrium for the game with full precision

PROPOSITION 3.2. *For any $c > 0$, $\lambda \in (0,1)$ there exists an equilibrium with full separation of types. The equilibrium level of effort satisfies (3.12) and it exceeds the social optimum.*

PROOF. We show that the strategies prescribed $a^*(\epsilon_i) = -\mu\epsilon_i$ and

$$d_i^*(a) = (1-\lambda)\nu a_i + \lambda\sum_{j\neq i}\nu a_j$$

are indeed an equilibrium. First, notice that the decision rule for the principal is

$$\begin{aligned} d_i(a) = (1-\lambda)E & [\theta_i] + \lambda\sum_{j\neq i}E\theta_j \\ &= (1-\lambda)E\left[a^*(\epsilon_i)+\epsilon_i \mid a_i\right] + \lambda\sum_{j\neq i}E\left[a^*(\epsilon_j)+\epsilon_j \mid a_j\right] \\ &= (1-\lambda)\left(a^{*-1}(a_i)+a_i\right) + \lambda\sum_{j\neq i}\left[a^{*-1}(a_j)+a_j\right] \\ &= (1-\lambda)\nu a_i + \lambda\sum_{j\neq i}\nu a_j \end{aligned}$$

Further, it is easy to see, that the distribution of actions is also symmetric around zero, so that $E\left[d_i(a) \mid a_i\right] = (1-\lambda)\nu a_i$. Hence, each agent minimizes

$$\begin{aligned} & E\left[(\theta_i - d_i(a))^2\right] + c(a) \\ s.t. E\left[d_i(a) \mid a_i\right] &= (1-\lambda)\nu a_i \end{aligned}$$

and the result follows. □

As we highlighted above, there exists multiple equilibria characterized by pooling of different types into the same decisions. In the reminder of this section we characterize them.

LEMMA 3.3. *Let $\{a^*(\cdot), d^*(\cdot)\}$ be an equilibrium for the game with perfect precision. If there exist $\epsilon < \epsilon'$, such that $a^*(\epsilon) = a^*(\epsilon')$, then, for all $\tilde{\epsilon} \in (\epsilon, \epsilon')$ $a^*(\tilde{\epsilon}) = a^*(\epsilon)$*

PROOF. In the Appendix A.3 □

This lemma implies that the equilibrium level of effort can be partitioned by a sequence of cutoffs $\{a_k\}$, such that, either $a_k = a(\epsilon)$ for all $\epsilon \in (\epsilon_k, \epsilon_{k+1})$, or $a^*(\epsilon) \in (a_k, a_{k+1})$ for all $\epsilon \in (\epsilon_k, \epsilon_{k+1})$. Further, whenever there is full separation in a given range, the equilibrium strategies must follow those in the previous proposition $a^*(\epsilon) = a_k - \mu(\epsilon - \epsilon_k)$. In particular, this lemma rules out equilibria involving not connected sets of types choosing the same action.

This results are important as a way to simplify the analysis of the optimal information structure. Indeed, if a set of types cannot be induced to choose a given action under some equilibrium of the full information system, actions cannot be supported as

LEMMA 3.4. *Under full precision, any equilibrium satisfies $d_k \leq d_{k'}$ whenever $a_k > a_{k'}$.*

PROOF. In the Appendix .3. □

3.4. Efficiency

In this section, we compare the efficiency properties of the equilibrium with full precision with those of equilibria in the game generated by partitions with pools. In particular, we show that for any possible environment (c and λ), and for any possible equilibrium played within the game generated by the finest partition, there is a coarser partition which attains a weakly higher payoff. To show this we construct a partition involving pooling in low types (that is, types close enough to zero). This information structure softens the constraints imposed in equilibrium under full precision, reducing the distortion in effort for all types, while incurring in information losses only in the selected interval. Thus, the principal is made better-off.

DEFINITION 3.5. An information structure presents pooling if there exists some $B_k \subset A$ with $\int_{B_i} dF(\epsilon) > 0$ and $m(a_i) = m(a_i')$ for all $a_i, a_i' \in B_k$

By Lemma 3.3, we will concentrate on information structures which can be characterized by a sequence of cutoffs. So that sets B_k can always be written as

$$B_k = [a_k, a_{k+1}]$$

PROPOSITION 3.6. *Suppose that* $\lambda c > 1$. *Then, for any distribution* F *and a* $\hat{\lambda}$, *such that for any* $\lambda \geq \hat{\lambda}$ *there exists an information structure* m' *with pooling with an equilibrium* $\{a'_k\}_{k=1,..K}$, *that is (weakly) more efficient than the information structure with full precision* m.

PROOF. In the Appendix A.3 □

Notice that this result does not hinge on any direct cost associated with a better information structure. It shows that even if information was completely cheap, not all the firms would buy a perfect system. In our constructive proof, we show that for small enough shocks, the gains accruing for more informed decisions are overwhelmed by the inefficient activities undertaken by the local units when they try to communicate that information. Hence, it is optimal for the HQ to pool those small types, forgo some precision in their decisions and reduce inefficient activities.

In what follows, we perform comparative statics to understand better which firms will adopt more precise information technologies according to our model

3.5. Comparative Statics

Which firms gain more from implementing information technologies to control the activities down the hierarchy? To answer this question we first need to obtain some characterization of the optimal information structure. Notice that Proposition 5 only tells us that if the need of coordination is sufficiently big firms are better off by acquiring some inefficient information technology. We do not know however, how inefficient this would be. To perform comparative statistics, we need to assume that the shock is uniform.

ASSUMPTION 3.1. ϵ_i *is uniformly distributed*

PROPOSITION 3.7. *In the optimal partition, there exists a thresold* $\epsilon^* \leq \bar{\epsilon}$ *such that there is perfect precision if* $|a| \geq a(\epsilon^*)$ *and (partial) pooling if* $|a| < a(\epsilon^*)$

PROOF. In the Appendix A.3 □

The intuition is similar to that of Proposition 3.6. If the principal prefers pooling at a given type, it must be that the improvement in effort allocation overcomes the decrease in the precision of the decision at those types. The improvement in the allocation is increasing in the number of types above, so that lower types are more attractive.

Our model is, therefore, able to predict that HQ will receive more precise information when adaptation is more important. Senior management should be called upon only on exceptional cases. This organizational design is usually referred as ”management-by-exception” and it has been discussed extensively in the literature.

PROPOSITION 3.8. *ϵ^* is increasing in the need of coordination λ and the support of the distribution $\bar{\epsilon}$*

PROOF. In the Appendix A.3 □

Increasing the need of coordination (weakly) reduces the optimal amount of information transmitted. This is because lower need for adaptation leads to less sensitivity of the decision of the HQ to information, more importance of local responses and lower value of information. Notice, that we can also see this result as a span-of-control type of result. As N grows, the amount of information optimally acquired about any of the agents decreases. The reason is not cognitive limitations but increasing agency costs. Furthermore, we see the solution as being akin to a management-by-exception type of organization, where only big problems are solved up in the hierarchy.

3.6. Empirical Relevance

There is a growing literature that studies the relation between organization and information technologies. Some examples are Bresnahan et al. (2002) and Bloom et al. (2012), where they study the effects of IT adoption on organizational design. Our focus is on the effect that different designs (and more broadly environment conditions that lead to some particular designs) have on the IT adoption by the firm.

First, we have shown that treating the IT as exogenous may bias significantly the results obtained in empirical applications. Increasing the acquisition of information for decision-making may result in worse incentives for lower-level units and, therefore, harm overall performance. If those firms who do not acquire better IT are those with more severe agency problems, an empirical analysis that takes IT as exogenous would conclude wrongly that IT improves performance. Interestingly, IT acquisition has been found to yield significant improvements in productivity only for those firms who perform organizational change, where these potential agency problems may be solved using other tools.

Second, we provide an explanation for the heterogeneity in the level of adoption of IT at the firm level. Firms with higher need for coordination would try to acquire information from outside the firm directly[10] and would be less willing to invest in IT systems that acquire information from inside[11]. Since these software had been consistently shown to be correlated with productivity gains (Aral et al.

[10]Chandler (1969) relates the way in which A. Sloan managed to get this information by driving to meet his dealers directly and getting data from the office of car registrations.

[11]Software like Enterprise Resource Planning (ERP), Supply Chain Management (SCM) and Customer Relationship Management (CRM).

(2007)), it was not clear why firms in different industries adopt them in an unequal way. In this sense, our chapter provides a simple explanation in terms of organizational design of the firm. Whenever, multi-divisional firms require more coordination or they confront a more uncertain environment, we expect them to have less incentives to acquire better internal information technologies.

Some of the qualitative results of this Chapter can be used to interpret the organizational architecture implemented by ABB, one of the most successful and complex firms in the world. ABB implemented a ”three-strikes-and-out” policy for conflict resolution (Roberts (2004)) Two managers who disagreed could take their issues to a higher level twice but they would be replaced after the third time. The idea is that only big issues should be dealt higher in the hierarchy and the probability of many of those between two units was really small. It is a general feature of multi-divisional organizations to rely on the hierarchy for conflict resolution only on very exceptional cases, as pointed out by Eccles (1985).

3.7. Conclusions

In this chapter, we tried to study the relationship between agency problems and Information Technologies in multi-divisional firms. Although the empirical literature has identified substantial efficiency gains for the acquiring IT firms (Aral et al. (2007)) there is large heterogeneity in the level of adoption across firms. In our model,even if IT's are useful to improve decision-making, they may create perverse incentives throughout the organization, leading to some firms to reduce their level of adoption. We have shown that the result depends crucially on the ability of the different units to signal credibly their need for adaptation and the inefficient activities this possibility generates. Our model makes testable predictions on the effect of span-of-control, need for coordination and the variability of the environment in the level of information technology adoption of the firm.

As for future research, it would be interesting to explore a model where communication could take place both through cheap talk and signalling activities. Cheap talk would increase the quality of information transmission for a fixed level of precision of the information structure but would also imply lower incentives for signalling through costly actions, and, therefore, the total effect on acquisition of information is ambiguous.

APPENDIX A

Appendix

A.1. Proofs of Chapter 1

PROPOSITION A.1. *A Stationary Markov Perfect Equilibrium Exists*

PROOF. If the game has a recursive formulation, then standard arguments show that such an Equilibrium Exists. The only complication is that the Value functions are not continuous (around 0). Nevertheless, all that is required for the recursive formulation to be well-defined is that Value Functions are Monotone. This is proved in a series of steps. First, in order to economize on space, I'll denote by p_μ is the probability with which the agent tells the truth given that he has received no signal.

CLAIM A.1. *Let* $\mu > 0$. $V(\mu) > V(\phi_A(\mu))$

PROOF. First, assume that $\sigma > 0$. Then, for a contradiction assume there exists some μ for which $V(\mu) \leq V(\phi_A(\mu))$. Notice that

$$\pi - c + \delta\left[\frac{p_\mu}{1+p_\mu}V(0) + \frac{1}{1+p_\mu}V(\mu)\right] = \pi_\mu + \delta V(\phi_A(\mu))$$

so that

$$\begin{aligned}\pi - c + \delta\left[\frac{p_\mu}{1+p_\mu}V(0) + \frac{1}{1+p_\mu}V(\mu)\right] &\geq \pi_\mu + \delta V(\mu)\\ \pi - c + \delta\frac{p_\mu}{1+p_\mu}V(0) &\geq \pi_\mu + \delta\frac{p_\mu}{1+p_\mu}V(\mu)\end{aligned}$$

which is true only for $\mu = 0$.

Assume now that $\sigma = 0$. Since

$$\begin{aligned}V(\mu) &= \pi_\mu + \delta\left[\frac{1}{2}(1+p_\mu)V(\phi_A(\mu)) + \frac{1}{2}(1-p_\mu)V(\phi_B(\mu))\right]\\ V(\mu) &\geq \frac{2\pi_\mu + \delta(1-p_\mu)V(\phi_B(\mu))}{2-\delta(1+p_\mu)}\end{aligned}$$

Notice that if $V(\phi_B(\mu)) < V(\mu)$, we have that $V(\phi_B(\mu)) \leq V(\phi_A(\mu))$ and we know that $V(\phi_B(\mu)) \leq V(\phi_A(\phi_B(\mu)))$ so substituting μ for $V(\phi_B(\mu))$ for a number of times, we will get as close as we want to $\mu = 1$ and we have a contradiction since $\lim_{\mu\to 1} V(\mu) = V(1) = \frac{\pi}{1-\delta} \geq V(\mu)$. So then assume that $V(\phi_B(\mu)) \geq V(\mu)$, then

$$V(\mu) \geq \frac{\pi_\mu}{1-\delta} = \frac{\pi}{1-\delta} - \frac{2q-1}{2(1-\delta)}(1-\mu)\bar{F}(x(\mu))$$

To see that this cannot be true if $\sigma = 0$, notice that $\bar{F}(x(\mu)) > \frac{1}{2}$ (to see why, just consider the simplified problem of the agent who is not allowed to go below the cutoff of the principal; maybe add a restriction of F so that this holds). It is clear then that

$$\begin{aligned} V(\mu) &\leq \mu\frac{\pi}{1-\delta} + (1-\mu)\left[\frac{\pi}{(1-\delta)} + \frac{\pi - c}{2(1-\delta)}\right] \\ &= \frac{\pi}{1-\delta} - (1-\mu)\frac{c}{2(1-\delta)} \end{aligned}$$

Hence a necessary condition is that

$$\begin{aligned} \frac{2q-1}{2(1-\delta)}(1-\mu)\bar{F}(x(\mu)) &\geq (1-\mu)\frac{c}{2(1-\delta)} \\ (2q-1)\bar{F}(x(\mu)) &\geq c \end{aligned}$$

but Assumption 1 guarantees that there exists $x_1 > 0$ for which $c = \bar{F}(x_1)(2q-1)$ and we know that this is the static equilibrium, where the value for the principal is much lower. Since for x higher than this, $c > (2q-1)\bar{F}(x(\mu))$ we reach a contradiction and $V(\mu) > V(\phi_A(\mu))$ □

CLAIM A.2. *V is increasing*

PROOF. For all μ such that that $\sigma = 0$ we have that

$$V(\mu) = q - (1 - \frac{p_\mu}{2})(2q-1) + \delta\left[\frac{1+p_\mu}{2}V(\phi_A(\mu)) + \frac{1-p_\mu}{2}V(\phi_B(\mu))\right]$$

where $p_\mu = (1-\mu)\bar{F}(x(\mu))$.

Since $\sigma > 0$, V increases if p_μ decreases. We can write $V(\mu)$ as

$$\begin{aligned} V(\mu) &= (1-\frac{1}{2}p_\mu)q + \frac{1}{2}p_\mu(1-q) + \delta\left[\frac{1}{2}(1+p_\mu)V(\phi_A(\mu)) + \frac{1}{2}(1-p_\mu)V(\phi_B(\mu))\right] \\ &= q - \frac{1}{2}(1+p_\mu)c + \delta\left[\frac{1}{2}p_\mu V(0) + \frac{1}{2}V(\mu) + \frac{1}{2}(1-p_\mu)V(\phi_B(\mu))\right] \end{aligned}$$

Now, suppose that there exists two different beliefs such that $\mu_H > \mu_L$ and $V(\mu_H) = V(\mu_L) = V$

$$\begin{aligned} V &= q - \frac{1}{2}(1+p_H)c + \delta\left[\frac{1}{2}p_H V(0) + \frac{1}{2}V + \frac{1}{2}(1-p_H)V(\phi_B(\mu_H))\right] \\ &= q - \frac{1}{2}(1+p_L)c + \delta\left[\frac{1}{2}p_L V(0) + \frac{1}{2}V + \frac{1}{2}(1-p_L)V(\phi_B(\mu_L))\right] \end{aligned}$$

Hence, we need that

$$\frac{1}{\delta}(p_H - p_L)c = (p_L - p_H)V(0) + (1-p_L)V(\phi_B(\mu_L)) - (1-p_H)V(\phi_B(\mu_H))$$

Suppose first that $p_H \geq p_L$. This requires that $V(\phi_B(\mu_L)) \geq V(\phi_B(\mu_H))$, where $\phi_B(\mu_H) > \phi_B(\mu_L)$. Now, notice that this would lead to a decreasing sequence of $\{\mu_{ki}; \mu_{ki} > \mu_{k-1i}\}_k^{i=1,2}$ but we know that $V(\mu)$ is strictly increasing when $\mu \to 1$. Hence, a contradiction. Suppose then that $p_H < p_L$. Notice that

$$q - c + \delta \left[\frac{p_H}{1+p_H} V(0) + \frac{1}{1+p_H} V \right] = \frac{q(1-p_H)+1}{1+p_H} + \delta V(\phi_A(\mu_H))$$

and

$$q - c + \delta \left[\frac{p_L}{1+p_L} V(0) + \frac{1}{1+p_L} V \right] = \frac{q(1-p_L)+1}{1+p_L} + \delta V(\phi_A(\mu_L))$$

so that

$$\delta \left[V - V(0) \right] \frac{p_L - p_H}{(1+p_H)(1+p_L)} = \frac{2q(p_L - p_H)}{(1+p_H)(1+p_L)} + \left[V(\phi_A(\mu_H)) - V(\phi_A(\mu_L) \right]$$

Since $V - V(0) \geq \lim_{\mu \to 0} V(\mu) = \frac{\epsilon}{1-\delta}$, for δ large enough, this requires that

$$V(\phi_A(\mu_H)) - V(\phi_A(\mu_L)) < 0$$

which again will not be true since $V(\mu)$ is strictly increasing for μ small enough. □

CLAIM A.3. *If $\sigma(\mu) > 0$ then U is non-decreasing and $(1-\mu)\bar{F}(x(\mu))$ is non-increasing.*

PROOF. Notice that if σ is weakly decreasing, the value of the agent is monotone. This follows from a standard choice argument, since the agent is free to choose his report and facing a lower monitoring intensity yields higher payoffs.

Now, for every μ such that $\sigma(\mu) > 0$ assume that V is non decreasing. There exists an open set of such μ by Assumption 2. From this it is easy to see that $(1-\mu)\bar{F}(x(\mu))$ must be non-increasing if σ is non-increasing. To see why, assume for a contradiction that $(1-\mu)\bar{F}(x(\mu))$ is increasing in some open interval (μ_0, μ_1). The principal must be indifferent between monitoring and not monitoring at every μ such that $\sigma(\mu) > 0$. Therefore

$$V_{A,1}(\mu) = V_{A,0}(\mu) \forall \mu \in (\mu_0, \mu_1)$$

but then the principal's value at the beginning of the period could be written as

$$V(\mu) = p_\mu q + (1-p_\mu)\pi_\mu + \delta \left[p_\mu V(\phi_B(\mu)) + (1-p_\mu) V_{A,0}(\mu) \right]$$

But $(1-p_\mu)$ and π_μ are increasing in μ so that either V is decreasing at μ or $V(\phi_B(\mu))$ is decreasing at μ. Finally notice that since at higher μ the agent lies more often, a report of B is more informative, thereby contradicting the monotonicity of V. □

CLAIM A.4. *Every equilibrium is monotone.*

PROOF. To see this notice that in non-monotone equilibiria p_μ is increasing and σ is increasing (at least in some open interval). But then in that open interval U must be decreasing. But in such a case (write the expression down)

$$U(\phi_B(\mu)) < U(\phi_A(\mu))$$

But then there exists a sequence of values for agents whose belief converges to zero that is increasing. But the limit of all these sequences is the infimum of all the values that are consistent with the principal randomizing. Therefore we have a subsequence converging from below to the infimum of all the equilibrium values. This is a contradiction. □

□

PROOF OF PROPOSITION 1.2. It is easy to see that the agent lies with positive probability for every μ. If this were not the case, the principal will not monitor when he holds that belief and, therefore, the agent's value of lying today is strictly positive. Since the agent does not lie in the prescribed equilibrium, both the strategic and the truthful type produce the same distribution of recommendations and the principal does not update his belief. Therefore, the future value does not depend on the recommendation. This contradicts the fact that the agent does not lie. Therefore, the biased agent lies for every μ and recommends A with higher probability than the truthful type and the belief follows a supermartingale. Since the principal's value of monitoring is bounded away from zero whenever the type of the agent is low enough, in any equilibrium the agent gets caught with probability 1. □

Before getting in the details of the Proof of Proposition 1.3, I sketch its main steps. Let $\Psi(\mu) = \mathbb{E}[V(\phi(\mu) \mid 1)] - \mathbb{E}[V(\phi(\mu) \mid 0)]$ be the dynamic value of the information. The proof shows first that if such value is non-negative for every μ, any Markov Perfect Equilibrium requires the agent to tell the truth with sufficiently high probability. From this observation, it follows that $V(\mu)$ is bounded below for $\mu > 0$ and away from $V(0)$. Finally, the fact that the agent tells the truth sufficiently often and that $V(\mu)$ has bounded variation, implies that $\Psi(\mu) < 0$ for μ small enough, thus contradicting the hypothesis. Then I show that if $\Psi(\mu) < 0$ in some interval (a, b), it must be that $a = 0$. This implies then that $\sigma^l(\mu) \leq \sigma^s(\mu)$ in such interval. This establishes Proposition 1.3.

Using this result, one can show that if $c > \bar{c}(\mu, q)$ then $\sigma^s(\mu_l) = 0$ and thus $\sigma^l(\mu) \leq \sigma^s(\mu)$ for all μ and $U^l(\mu) \geq U^s(\mu)$. This establishes Proposition 1.4.

PROOF OF PROPOSITION 1.3. Assume that $\Psi(\mu) > 0$ in any interval $(0, x)$. Clearly, in any MPE such that the principal monitors and

$\mu > 0$

$$(1-\mu)(q - \mathbb{E}(v(\mu))) + \Psi(\mu) = c$$

since for μ^*,

$$(1-\mu^*)(q-(1-q))\frac{1}{2} = c$$

if $\Psi(\mu) > 0$,

$$(1-\mu)(q - \mathbb{E}(v(\mu))) < (1-\mu^*)(q-(1-q))\frac{1}{2}$$

for every $\mu > 0$, so that

$$\begin{aligned} q - \mathbb{E}(v(\mu)) &\leq (1-\mu^*)(q - \frac{1}{2}) \\ \mathbb{E}(v(\mu)) &\geq q\mu^* + \frac{1}{2}(1-\mu^*) \end{aligned}$$

That is, the strategic agent reports B with probability higher than $\mu^* > 0$. The value for the principal is thus

$$V(\mu) \geq V(0) + \frac{\mu^*}{2-\delta\mu^*}c$$

But notice that $V(\mu) \leq (1-\mu)\lim_{x\to 0} V(x) + \mu V(1)$ by standard arguments. Thus

$$\begin{aligned} V(\mu) - \lim_{x\to 0} V(x) &\leq \mu\left[V(1) - \lim_{x\to 0} V(x)\right] \\ &\leq \mu\left[V(1) - V(0) - \frac{\mu^*}{2-\delta\mu^*}c\right] \\ &= \frac{\mu}{1-\delta}c\left[\frac{2-(1-\delta)\mu^*}{2-\delta\mu^*}\right] \end{aligned}$$

and that

$$\begin{aligned} \phi_A(\mu) &= \mu\frac{1}{1+(1-\mu)\bar{F}(x^l(\mu))} \\ &\leq \mu\frac{1}{1+(1-\mu)(1-\mu^*)} \end{aligned}$$

Let $\mu_l \in (0,1)$ satisfy

$$\frac{\mu_l}{1-\delta}\left[\frac{2-(1-\delta)\mu^*}{2-\delta\mu^*}\right]\frac{(1-\mu_l)(1-\mu^*)}{1+(1-\mu_l)(1-\mu^*)} = \frac{\mu^*}{2-\delta\mu^*}$$

which does not depend on q or c except for μ^*, if such a value exists or else $\mu_l = 1$. Notice that $\mu_l \leq 0$ is not a solution. Recall that

$$\Psi(\mu) = \frac{1}{1+(1-\mu)\bar{F}(x^l(\mu))}\left[V(\mu) - V(\phi_A(\mu))\right] + \frac{(1-\mu)\bar{F}(x^l(\mu))}{1+(1-\mu)\bar{F}(x^l(\mu))}\left[V(0) - V(\phi_A(\mu))\right]$$

but for $\mu < \mu_l$, $V(\mu) - V(\phi_A(\mu) < \frac{\mu^*}{2-\delta\mu^*}c$ and hence

$$\begin{aligned}\Psi(\mu) &< \frac{1}{1+(1-\mu)\bar{F}(x^l(\mu))}\frac{\mu^*}{2-\delta\mu^*}c - \frac{(1-\mu)\bar{F}(x^l(\mu))}{1+(1-\mu)\bar{F}(x^l(\mu))}\frac{\mu^*}{2-\delta\mu^*}c \\ &< \frac{(1-\mu)\bar{F}(x^l(\mu))-1}{1+(1-\mu)\bar{F}(x^l(\mu))}\frac{\mu^*}{2-\delta\mu^*}c < 0\end{aligned}$$

by construction. Thus $\Psi(\mu) < 0$ in every interval $(0, x)$ with $\mu < \mu_l$. Therefore, for the long-run principal to randomize, the agent must lie with higher probability at every $\mu < \mu_l$ and $x^s(\mu) < x^l(\mu)$. The fact that $\sigma^s(\mu_t) > \sigma^l(\mu_t)$ is shown next under the assumption that the value of the agent is also higher. If the value where lower, the result comes directly. □

PROOF OF PROPOSITION 1.4. Since the agent only cares about the monitoring intensity, the ranking in payoffs obtains if $\sigma^l(\mu) \leq \sigma^s(\mu)$. If $\mu_l = 1$ the result is inmediate. If not, let $\mu^* = \mu_l$ and notice that since that solution was independent of q and c, I can pick μ^* freely. Thus, it is clear that for $\mu > \mu^*$ there is no monitoring in either game. For $\mu < \mu^*$ the patient principal monitors less often if the agent cheats at most as much. Now, in equilibrium, the agent will increase his cheating given that the principal has less incentives to monitor but he will also get a higher payoff.

Now, I show that it cannot exist an equilibrium satisfying, $\sigma^s(\mu) \leq \sigma^l(\mu)$ for everty μ, there must exist another equilibrium $(\tilde{\sigma}, \tilde{x})$ for the game with a short-run player such that $\tilde{\sigma}^s(\mu) \geq \sigma^l(\mu)$. First notice that, given the previous discussion, if $\sigma^s(\mu) \leq \sigma^l(\mu)$ it must be true that $x^s(\mu) \geq x^l(\mu)$, for otherwise the incentives of the principal to monitor at some belief must be lower. The following argument shows that this cannot be part of an equilibrium. Notice that

$$\begin{aligned}x^l(\mu)(1-2q+\lambda)(1-\sigma^l) &= U^l(\phi_B(\mu)) - \mathbb{E}_A(U^l(\mu)) \\ x^s(\mu)(1-2q+\lambda)(1-\sigma^s) &= U^s(\phi_B(\mu)) - \mathbb{E}_A(U^s(\mu))\end{aligned}$$

Given that $x^l(\mu) \leq x^s(\mu)$, for every μ,

$$U^l\left(\phi_B^l(\mu)\right) - \mathbb{E}_A(U^l(\mu)) \leq U^s\left(\phi_B^s(\mu)\right) - \mathbb{E}_A(U^s(\mu))$$

Let $\Delta(\mathbb{E}) = \mathbb{E}_A(U^l(\mu)) - \mathbb{E}_A(U^s(\mu))$

$$\begin{aligned}\Delta(\mathbb{E}) &= (1-\sigma^l)U^l\left(\phi_A^l(\mu)\right) - (1-\sigma^s)U^s\left(\phi_A^s(\mu)\right) + \left(\sigma^l - \sigma^s\right)U(0) \\ &= (1-\sigma^l)(U^l\left(\phi_A^l(\mu)\right) - U^s\left(\phi_A^s(\mu)\right)) + \left(\sigma^l - \sigma^s\right)\left(U(0) - U^s\left(\phi_A^s(\mu)\right)\right)\end{aligned}$$

Clearly $\left(\sigma^l - \sigma^s\right)\left(U(0) - U^s\left(\phi_A^s(\mu)\right)\right) < 0$. Thus we have

$$U^l\left(\phi_B^l(\mu)\right) - U^s\left(\phi_B^s(\mu)\right) \leq (1-\sigma^l)U^l\left(\phi_A^l(\mu)\right) - U^s\left(\phi_A^s(\mu)\right))$$

or

$$U^l\left(\phi_B^l(\mu)\right) - (1-\sigma^l)U^l\left(\phi_A^l(\mu)\right) \leq U^s\left(\phi_B^s(\mu)\right) - (1-\sigma^l)U^s\left(\phi_A^s(\mu)\right)$$

But notice that $\phi_A^l(\mu) < \phi_A^s(\mu)$ while $\phi_B^l(\mu) > \phi_B^s(\mu)$ since the agent lies more against a long run player. Thus the function U^s is more "steep" than the function U^l for every μ. Since $U(1)$ is the same for both, it must be that $U^l(\mu) > U^s(\mu)$ for almost every μ. Thus, it cannot be that $\sigma^s(\mu) \leq \sigma^l(\mu)$ for every μ.

Notice that if $U^l(\mu) \geq U^s(\mu)$ for every μ, if $\sigma^s(\mu) < \sigma^l(\mu)$ for some $\mu < \mu^*$ it must be that $x^s(\mu) > x^l(\mu)$. But in that case, the agent playing against a long-run principal faces a higher likelihood and a higher cost in case of being caught and, therefore, cannot be indifferent. Therefore $x^s(\mu) \leq x^l(\mu)$ and $\sigma^s(\mu) \geq \sigma^l(\mu)$.

□

PROOF OF PROPOSITION 1.5. The intuition for this result is the standard Stackelberg-Cournot logic. On the margin, the principal is willing tocommit to increase her monitoring intensity since this encourages truthfulreporting. Notice that, in equilibrium, and conditional on the agent telling the truth or not, if the principal monitors at all she is indifferent between doing so or not.

$$V_{A,1}(\mu) = V_{A,0}(\mu)$$

With one-period commitment, however, the principal can affect the probability with which the agent tells the truth. She chooses σ to solve

$$\max_{\sigma} \frac{1}{2}(1 + (1-\mu)F(\bar{x}(\mu;\sigma))V_A(\mu) + \frac{1}{2}(1 - (1-\mu)F(\bar{x}(\mu;\sigma))V_B(\mu)$$

At the equilibrium σ and if $\sigma > 0$ we have

$$\max_{\sigma} \frac{1}{2}(1 + (1-\mu)F(\bar{x}(\mu;\sigma))V_{A,1}(\mu) + \frac{1}{2}(1 - (1-\mu)F(\bar{x}(\mu;\sigma))V_B(\mu)$$

which can be rewritten as

$$\max_{\sigma} \frac{1}{2}\left[V_{A,1}(\mu) + V_B(\mu)\right] - (1-\mu)F(\bar{x}(\mu;\sigma)\left[V_B(\mu) - V_{A,1}(\mu)\right]$$

so that clearly, by increasing σ the principal can increase his payoff.

□

PROOF OF PROPOSITION 1.6. The argument proceeds in two steps. I first show that there cannot be a fixed time $t^* > 1$ such that, the type of the agent has been revealed for sure, and the principal attains V^g if the agent is good. In the second step I extend the result to stochastic time limits.

To see the first claim, notice that at time $t^* - 1$, and independently of the way in which the type is revealed, a strategic agent who has not been revealed must choose action A independently of his current information. In particular, it must be the case that

$$\begin{aligned} x_l\left[(1-q+\lambda)(1-\bar{\sigma})+\bar{\sigma}q\right]+\delta U^b &\geq x_l q+\delta U^g \\ &= x_l q+\frac{\delta}{1-\delta}\bar{x}(\frac{1}{2}+\lambda) \end{aligned}$$

while $U^b=\frac{q}{1-\delta}$. Hence

$$x_l(1-2q+\lambda)(1-\bar{\sigma})+\delta\frac{q}{1-\delta}\geq\frac{\delta}{1-\delta}\bar{x}(\frac{1}{2}+\lambda) \tag{A.1}$$

since $\frac{1}{2}+\lambda>q$ by assumption, for every $\bar{\sigma}$, there is a δ^*, such that if $\sigma>\delta^*$, the condition is violated.

Now, for the second claim, notice that, if the IC constraint is to be satisfied, the maximum intensity after every history h is

$$x(h)(1-2q+\lambda)(1-\sigma(h))+\delta\left[(1-\sigma(h))U(h\sqcup a)+\sigma(h)U^b\right]\geq\delta\bar{x}U(h\sqcup b)$$

Since $U^b<U(h\sqcup a)\leq U(h\sqcup b)$, this gives an upper bound on $\sigma(h)$ for every $x(h)$. Hence, let $\bar{\sigma}(h)$ be such an upper bound. Since $\bar{\sigma}(h)\in(0,1)$ and $x\in[x_l,x_h]$, both variables define a probability measure over histories. Let

$$\Omega^t=\left\{h_t\in\mathcal{H}^t \text{ s.t. } \mu\in\{0,1\}\right\}$$

Let $p(\Omega^t)$ be the implied probability of such event. All we have to show is that

$$C=(1-\delta)\sum_{h^t\in\Omega^t}p(h^t)\sum_{t=0}\delta^t\sigma(h^t)c>0 \tag{A.2}$$

since in that event $V^g<\frac{q}{1-\delta}-\epsilon$ for some $\epsilon>0$, or $V^b<\frac{q-c}{1-\delta}$ since

$$(1-\delta)\sum_{h^t\in\Omega^t}p(h^t)\sum_{t=0}\delta^t\bar{F}(x(h^t))(2q-1)>0$$

. We have

$$\begin{aligned} p(\Omega^t)-p(\Omega^{t-1}) &= \left[\frac{1}{2}\int\sigma(h^t)F(x(h^t)dp(h^t\mid h^{t-1}\notin\Omega^{t-1})\right]\left[1-p(\Omega^t)\right] \\ &\leq \frac{1}{2}\left[1-p(\Omega^t)\right]\sup_{h^t\in\Psi(\mathcal{H}^{t-1}\setminus\Omega^{t-1})}\bar{\sigma}(h^t)F(x(h^t)) \end{aligned}$$

But because of the monotonicity property in the value function[1].

$$\sup_{h^t\in\Psi(\mathcal{H}^{t-1}\setminus\Omega^{t-1})}\bar{\sigma}(h^t)F(x(h^t))\leq\bar{\sigma}(h^*)F(x(h^*))$$

[1]In the optimal solution to the Full Commitment problem the value function is monotone as long as the value for the agent is monotone. This value, in turn, is monotone as long as the monitoring intensity of the agent is non-increasing. A simple local indifference argument shows that this must be true in the FC case.

where at h^*, $U(h^* \sqcup a) = U(h^* \sqcup b)$, and hence using (A.1)

$$x(h^*) = \delta \frac{\sigma(h^*)(U(h \sqcup r) - U^b)}{(1 - 2q + \lambda)(1 - \sigma(h^*))}$$

Clearly this condition imposes a bound on $F(x(h))$ for every $\sigma(h)$. Let that bound be $\bar{F}(\sigma)$. Notice that the bound depends both explicitly and implicitly on δ through the continuation values of the agent. To further simplify the problem, let us fix the average expected continuation monitoring intensity from the point of view of the agent to be some $\sigma' < 1$ so that

$$U(h \sqcup r) - U^b \geq \frac{1}{1 - \delta\sigma'} \left[1 - 2q + \lambda\right] \bar{x} \frac{1}{2}$$

And thus

$$x(h^*) \geq \delta \frac{\sigma}{1 - \sigma} \frac{\bar{x}}{1 - \delta\sigma'} \frac{1}{2}$$

As $\delta \to 1$

$$x(h^*) \geq \frac{\sigma \bar{x}}{2(1 - \sigma)\,(1 - \sigma')}$$

First assume that $\sigma \geq \frac{2(1-\sigma')}{1+2(1-\sigma')}$. In such a case, $F(x(h)) \leq \frac{1}{2}$. Thus $\bar{\sigma}(h^t)F(x(h^t)) < \frac{1}{2}$ for every h^t and

$$p(\Omega^t) - p(\Omega^{t-1}) < \frac{1}{4} \left[1 - p(\Omega^t)\right]$$

while $\sigma \geq \frac{2(1-\sigma')}{1+2(1-\sigma')} > 0$. Thus (A.2) becomes

$$\begin{aligned} C &\geq (1 - \delta) \sum_t \left(\frac{3\delta}{4}\right)^t c \\ &= (1 - \delta) \frac{4}{4 - 3\delta} c \end{aligned}$$

Now if $\sigma < \frac{2(1-\sigma')}{1+2(1-\sigma')}$, since σ was the upper-bound of all possible σ we have that $\sigma \geq \sigma'$, and so $\sigma \leq 2(1 - \sigma)^2$, so that $\sigma \leq \frac{1}{2}$

$$p(\Omega^t) - p(\Omega^{t-1}) \leq \frac{1}{4} \left[1 - p(\Omega^t)\right]$$

and hence (A.2) does also hold. □

A.2. Proofs of Chapter 2

In this Appendix we show first that there exists a Linear Equilibrium under Perfect Information Transmission as long as the Economy is large enough compared with the maximum degree of a given player. The rest of the Appendix contains omitted proofs

LEMMA A.2. *Assume that the network is balanced. Then, a Linear Equilibrium exists. The weight that a given player puts on his neighbor's signal is decreasing in the amount of information he has access to and increasing in the information this player provides and in the centrality measure of his neighbor.*

PROOF. The argument is standard. Assume everyone else follows a linear strategy, and let agent i have a neighborhood $N_i(g)$. He solves

$$\min E_i\left(\theta - a_i\right)^2 + \frac{1}{n-1}\sum_{j\neq i} E_i(a_i - a_j)$$

$$s.t.\ E_i(a_j) = \sum_{k\in N_{ij}^*} b_{jk}x_k + \sum_{k\in N_{i-j}} b_{jk}E_i(x_k)$$

First Order Condition is

$$a_i = \frac{1}{2}E_i(\theta) + \frac{1}{2(n-1)}\left\{\begin{array}{c} \sum_{j\in N_i(g)}\left[\sum_{k\in N_{ij}} b_{jk}x_k + E_i(\theta)\sum_{k\in N_{i-j}} b_{jk}\right] + \\ \sum_{j\notin N_i^*(g)}\left[\sum_{k\in N_{ij}} b_{jk}x_k + E_i(\theta)\sum_{k\in N_{j-i(g)}} b_{jk}\right]\end{array}\right\}$$

We can rewrite this expression as

$$a_i = \frac{1}{2}\left\{1 + \frac{1}{n-1}\sum_{j\neq i}\sum_{k\in N_{j-i}^*} b_{jk}\right\}E_i(\theta) + \frac{1}{2(n-1)}\sum_{j\neq i}\sum_{k\in N_{ij}^*} b_{jk}x_k$$

Since

$$E_i(\theta) = \sum_{k\in N_i^*(g)} \frac{\tau_k x_k}{\sum_{k\in N_i(g)}\tau_k + \tau_\sigma}$$

we can write

$$a_i = \sum_{k\in N_i^*(g)} b_{ik}x_k$$

where the vector b may be identified through matrix algebra.

As it is well known, if $\tau_\sigma = 0$, $\sum b_{ik} = 1$ for all $i \in N$. Clearly if $\tau_\sigma > 0$, and the network belongs to the core-periphery network (see Corollary 2) $b_{ik} = \frac{\tau_k}{\sum \tau_k + \tau_\sigma}$, $\sum b_{ik} < 1$, which is an upper bound for the other topologies. If the network structure is regular we have $b_{ik} = b$ for all $i, k, i \in N_i^*(k)$. Thus, letting m be the degree of the network we have

$$
\begin{aligned}
a_i &= mb \sum_{k \in N_i(g)} x_k \\
b &= \left[1 - \frac{K-M}{2(n-1)}\right]^{-1} \frac{1}{2} \frac{\tau}{m\tau + \tau_\sigma} \left[1 + \frac{M}{n-1}\right]
\end{aligned}
$$

where M is the number of links in the network and K is the number of links not in the network. □

PROOF OF PROPOSITION 2.4. Order them from 1 to n, so that $ii + 1 \in g^T$.If $i \in \{1, n\}$, the result follows because $\tau_3, \tau_{n-2} > 0$. Assume then that $N_i(g) = \{i-1, i+1\}$. I have to show that if i can communicate truthfully with $i+1$, then it cannot communicate truthfully with $i-1$. For a contradiction, assume it is not true. That is that i can communicate with both. Notice that it follows that their residual variances are the same. Assume that i can communicate with $i+1$. Clearly the sequence $\{1, 2, ..., i-1, i\}$ is a path that links i with 1. Notice that for 2 the result follows. By induction, assume that it holds for i, we show that it must hold for $i+1 \leq \frac{n+1}{2}$. For the rest of agents the prove is symmetric and thus omitted. Notice that $i+1$ if could communicate truthfully with i and with $i-1$, and since $i-1$ has lower ability to coordinate than $i+1$ and the same residual variance, so by proposition 4, $i-1$ should decrease her information acquisition (since all agents acquire some information this is feasible) and hence that could not be an equilibrium. □

PROOF OF PROPOSITION 2.5 . The strategic substitutability of the information acquisition effort between neighbors stems from the public good nature of the information. I'll show now the complementarity Clearly, since g^T is connected and balanced, it must be the case that

$$\sum_{i \in N_i^*(g)} \tau_i = \bar{\tau} > 0 \text{ for all } i \in N$$

Returns from information are
(A.3)

$$\frac{\partial}{\partial \tau_i} E\left[(a_i - \theta)^2 + \frac{\sum_{j \neq i}(a_i - a_j)^2}{n-1}\right] = \frac{\partial}{\partial \tau_i} \left\{ \begin{array}{c} E\left[(\sum_{j \in N_i^*(g^T)} b_{ij} x_{ij} - \theta)^2\right] + \\ \frac{1}{n-1} \sum\limits_{j \in N} E\left[(\sum_{l \in N_i^*(g)} b_{il} x_{il} - \sum_{l \in N_j^*(g)} b_{jl} x_{jl})^2\right] \end{array} \right\}$$

The first term in the last line of the equation measures the marginal value of information to predict the state of the world. Clearly, it is constant for all agents since all their actions are equally informed. It decreases in the effort of other players. The second term can be rewritten as

$$\begin{aligned} V_i &= \frac{1}{n-1}\sum_{j\in N} E\left[(\sum_{l\in N^*_{i-j}} b_{il}x_{il} - \sum_{l\in N^*_{j-i}} b_{jl}x_{jl})^2 \right] = \\ & \frac{1}{n-1}\sum_{j\in N} E\left[(1-\sum_{l\in N^*_{i-j}} b_{il})\theta^2 + (\sum_{l\in N^*_{j-i}} b_{il}\eta_{il} - \sum_{l\in N^*_{i-j}} b_{jl}\eta_{jl})^2 \right] \\ &= \frac{1-\sum_{l\in N^*_{i-j}} b_{il}}{\tau_\sigma} + \frac{1}{n-1}\sum_{j\in N} E\left[(\sum_{l\in N^*_{j-i}} b_{il}\eta_{il} - \sum_{l\in N^*_{i-j}} b_{jl}\eta_{jl})^2 \right] \end{aligned}$$

This term measures the amount of coordination losses brought about by the dispersion and incompleteness of information. In particular, the first term measures the value of the information about θ when trying to predict other people's actions -as in Hellwig and Veldkamp (2009)-. The second term measures the value of the reports of others when trying to forecast other players' actions. To see the monotonicity, notice that V_i increases in the effort of other players, since higher effort implies that learning the true value of θ becomes more informative about other players' actions. This implies that if i and j are not linked, their efforts are strategic complements. □

Proof of Proposition 2.11. There are two possibilities. First assume that no information is withheld. Then, assume for a contradiction that information revelation is perfect. Then, it must be the case that every two individuals obtain the same amount of information at time t_i^* where t_i^* is such that i takes the action. Notice that in a pure strategy equilibrium t_i^* is deterministic (in particular, it does not depend on the realizations of the signals). Clearly, 1 should leave in the same period as 2, since in the following period 1 will not receive new information[2]. However, at the period in which 2 leaves, if optimal, he shall get at least one more signal. Hence, 1 is always less informed than 2 and results in Proposition 3 apply.

It is also straightforward to see that every agent leaving at period $t = 1, 2...$ it is an equilibrium, provided sufficiently many signals are obtained. In particular $\bar{t}$ would be the minimum between the period at which the value of two more signals to the most central agent is lower than δ and $\frac{n+1}{2}$.

To conclude, we show that withholding of information does not change the results . First notice that withholding information to j for less than t_j^* periods is never optimal (i would just reduce his own influence on other players obtaining the same amount of information.)

[2]If the equilibrium involves mixed strategies, the strategy of player 1, conditional on observing that player 2 left is to leave in the following period, but no more information is revealed to him.

Now, suppose that agent i conceals his information until period t_j^* - that is, the period at which j leaves, and assume that i and j have access to the same information, then we show that $i-1$ must have access to less information than them. If i and j have access to the same information and $t = t_j^*$ they both leave. Then, if $i-1$ is to have the same amount of information as them, it must be that he receives a report of the same precision at period t_j^* (or later), and then leave. However, in the line, this requires that there exists another agent $i - t_j^*$ who originated that report and now gets to $i-1$. Now, if that is the case, then at period t_j^*, j must have received another report coming from agent $i-1-t_j^*$, and thus, j has access to more information than i. □

PROOF OF PROPOSITION 2.12 . The idea is similar as in Proposition 2.11. Suppose the result does not hold. Then, there exists $j \in N_i(g)$ such that i can communicate truthfully with j. We know that it must be the case that their residual variances are equal, and thus they have received the same number of signals. Since i's neighborhood is a subset of that of j, this can happen if and only if in the last round of communication whatever j learns also i does. Hence, it must be that j i) does not receive information that was not held by another agent in the neighborhood of i in the previous period and ii) decides not to leave until he gets to that stage. If δ is sufficiently large, he will leave before. If δ is sufficiently small, however, he will stay until all information is received. Since this must happen for all $j \in N_i(g)$ in order for i to communicate truthfully, it must be the case, that at time $t_i^* = t_j^*$ for all $j \in N_i(g)$, no new information arrives to the neighborhood of i so that all information must be aggregated before everyone leaves, thus establishing the claim. □

A.3. Proofs of Chapter 3

PROOF OF LEMMA 3.3. Suppose on the contrary that there exist three types $\epsilon_1 < \epsilon_2 < \epsilon_3$ such that $a(\epsilon_1) = a(\epsilon_3) = a$ and $a(\epsilon_2) = a' \neq a$ with $d \leq d'$. Revealed preference implies that

$$\begin{aligned}(\epsilon_1 + a - d)^2 - (\epsilon_1 + a' - d')^2 &\leq (\epsilon_2 + a - d)^2 - (\epsilon_2 + a' - d')^2 \\ (\epsilon_3 + a - d)^2 - (\epsilon_3 + a' - d')^2 &\leq (\epsilon_2 + a - d)^2 - (\epsilon_2 + a' - d')^2\end{aligned}$$

Since this must hold for all $\epsilon_3 \in$supp(d) it must also hold for the $\max\{\epsilon : d(a(\epsilon)) = d\}$. Since the decision is chosen optimally it must hold that $d < \epsilon_3 + a$, so the second line implies that $\epsilon_2 + a < d$, which implies that $\epsilon_1 + a < d$. Hence, both agents have aligned preferences and the reverse order is impossible. □

PROOF OF LEMMA 3.4. The proof is by induction, taking only the positive set of types - the other one is symmetric. It is well-known that in this class of partition equilibria, the boundaries of each interval must satisfy an indifference condition between the two limiting intervals. Since zero types are not biased, in any equilibrium they must be able to truthfully reveal their type. Let $0 \in (\epsilon_l, \epsilon_{l+1})$, so that $0 \in (a_l, a_{l+1})$. Let $\zeta_l = E\left[\epsilon \mid \epsilon \in (\epsilon_l, \epsilon_{l+1})\right]$. Notice that ϵ_{l+1} must satisfy

$$(\epsilon_{l+1} - \lambda\zeta_l)^2 = (\epsilon_{l+1} + a_{l+1} - \lambda(\zeta_{l+1} + a_{l+1})^2 + c(a_{l+1}) \tag{A.4}$$

which requires $\lambda a_{l+1} < \lambda\zeta_l - \lambda\zeta_{l+1}$, $a_{l+1} < 0$. But then either $(\epsilon_{l+1} - \lambda\zeta_l)^2 < (\epsilon_{l+1} + a_{l+1} - \lambda(\zeta_{l+1} + a_{l+1}))^2$, so that (A.4) cannot hold; or there exists a type $0 < \epsilon < \epsilon_{l+1}$, such that $\epsilon + a_{l+1} > \lambda(\zeta_{l+1} + a_{l+1})$, who prefers a_{l+1} to 0. Hence $\lambda(\zeta_{l+1} + a_{l+1}) = E(d_i(a_l)) > 0$.

Now, suppose that $E(d_i(a_{k+1})) > E(d_i(a_k))$, for all $k \geq l$. Then we show, that $E(d_i(a_{k+2})) > E(d_i(a_{k+1}))$. Notice for a contradiction, that by a similar argument if the relation were not to hold, it must be that $a_{k+2} < a_{k+1}$. Because ϵ_{K+1} is indifferent between a_{k+1} and a_k, it cannot be that he prefers a_{k+2} to one of them. In particular, he cannot prefer a_{k+2} to a_{k+1}. It follows that $c(a_{k+2}) > c(a_{k+1})$, so that $a_{k+2} < 0$. But then, either ϵ_{k+2} prefers a_{k+1}, since the action is cheaper and the decision is closer, or if the decision is further away, it must be that ϵ_{k+1} prefers a_{k+2} to a_{k+1}. □

PROOF OF PROPOSITION 3.6. First notice that if the equilibrium did not involve perfect separation of types, the principal can be made (weakly) better-off by replicating the equilibrium choosing to observe all the actions that belonged to the support of the equilibrium strategy and pool the rest. If the equilibrium entails full separation, consider the following perturbation.

$$\begin{aligned} E(a \quad | \quad & m(a)) = a \; if \; a \in A \setminus (-a', a') \\ = \quad & E(a \mid a \in (-a', 0)) \text{ if } a \in (-a', 0) \\ = \quad & E(a \mid a \in (0, a')) \text{ if } a \in (0, a') \end{aligned}$$

Notice that if $a_i \in (-a', 0)$, and if $a^*()$ is invertible

$$d_i(m(a_i), m(a_{-i})) = \chi_k + \lambda \sum_{i \neq j} E(\theta_{-i} \mid a_{-i})$$

where $\chi_k = \frac{1}{F(a^{*-1}(a')) - F(a^{*-1}(0))} \int_{a^{*-1}(a_k)}^{a^{*-1}(a_{k+1})} (\varepsilon + a(\epsilon))\, dF(\varepsilon)$. Suppose $\lambda c > 1$. Then we have that

$$a(\epsilon) = \begin{cases} a' + \mu^*(\epsilon - a^{-1}(a')) \text{ for almost all } \epsilon \in (-\bar{\epsilon}, a^{-1}(a')) \\ \min\{0, \alpha + \beta\epsilon\} \text{ for almost all } \epsilon \in (a^{-1}(a'), 0) \\ \max\{0, -\alpha + \beta\epsilon\} \text{ for almost all } \epsilon \in (0, a^{-1}(-a')) \\ a' + \mu^*(\epsilon - a^{-1}(-a')) \text{ otherwise} \end{cases}$$

with $\beta = -\frac{1}{1+c}$ and α uniquely determined by maximization conditions. Hence, the only choice is a'. Notice that for such a scheme to be optimal

$$|a'| < \left|\mu^* a^{-1}(a')\right| = -\mu^*\epsilon_1$$

Further, the agent with type ϵ_1 must be indifferent between pooling or revealing. Hence

$$\begin{aligned} c\left(a'\right)^2 + E\left[(\epsilon_1 + a' - d(a'))^2\right] &= c(\min\{0, \alpha+\beta\epsilon_1\})^2 + \\ &+ E\left[(\epsilon_1 + \min\{0, \alpha+\beta\epsilon_1\} - d(m(\min\{0, \alpha+\beta\epsilon_1\})))^2\right] \end{aligned}$$

equivalent to

$$\begin{aligned} c\left(a'\right)^2 + E\left[(\epsilon_1 + a' - d(a'))^2\right] &= c(\alpha+\beta\epsilon_1)^2 + \\ &+ E\left[(\epsilon_1 + \alpha + \beta\epsilon_1 - d(m(\alpha+\beta\epsilon_1)))^2\right] \end{aligned}$$

given $|a'| < -\mu^*\epsilon_1$, we need that there exists $(\epsilon_1, \delta(\epsilon_1))$, with $\epsilon_1 < 0$, $\delta > 0$, satisfying

$$\begin{aligned} c\left(-\mu^*\epsilon_1\right)^2 + E\left[(\epsilon_1 + \mu^*\epsilon_1 - d(\mu^*\epsilon_1))^2\right] &= c(\alpha+\beta\epsilon_1)^2 + \\ &+ E\left[(\epsilon_1 + \alpha + \beta\epsilon_1 - d(m(\alpha+\beta\epsilon_1)))^2\right] + \delta \end{aligned}$$

Using our bound for μ^* whenever $\lambda c > 1$. Righ-hand side becomes

$$\begin{aligned} RHS &= c\mu^{*2}\epsilon_1^2 + \epsilon_1^2(1+\mu^*)^2 + \\ &+ E(d(\mu^*\epsilon_1))^2 - 2\epsilon_1(1+\mu^*)E\left(d(\mu^*\epsilon_1)\right) \\ &= \epsilon_1^2(c\mu^{*2} + (1+\mu^*)^2\lambda^2) + \lambda^2 E(\frac{1}{N}\sum(\epsilon_j + a(\epsilon_j))^2) \end{aligned}$$

whereas Left-hand side becomes

$$\begin{aligned} LHS &= c\alpha^2 + c\beta^2\epsilon_1 + 2c\alpha\beta\epsilon_1^2 + E\left[(\alpha + (1+\beta)\epsilon_1 - d(m(\alpha+\beta\epsilon_1)))^2\right] \\ &= c\alpha^2 + 2c\alpha\beta\epsilon_1 + c\beta^2\epsilon_1^2 + (\alpha+(1+\beta)\epsilon_1)^2 + (1-\lambda)\chi_k + \lambda^2 E(\frac{1}{N-1}\sum(\epsilon_j + a(\epsilon_j))^2 \end{aligned}$$

So the last term cancels and

$$\epsilon_1^2(c\mu^{*2} + (1+\mu^*)^2\lambda^2) \geq c\alpha^2 + 2c\alpha\beta\epsilon_1 + c\beta^2\epsilon_1^2 + (\alpha+(1+\beta)\epsilon_1)^2 + [(1-\lambda)\chi_k]^2$$

Both terms are zero at zero and for ϵ_1 small enough, the derivative of the LHS is the sum of a negative term and a term proportional to the derivative of the square of the expectation of the type within some range with respect to the upper bound. This derivative is bounded by the bound. Hence, as long as

$$c(\mu^{*2} - \beta^2) + (1+\mu^*)^2\lambda^2 > (1+\beta)^2 + (1-\lambda)^2$$

which happens for λ big enough, $c\lambda > 1$, you can take $a(\epsilon_1) < \mu\epsilon_1$ and that is independent of $F(\epsilon_1)$. Hence, as long as ϵ_1 is small enough, if there exists a $\mu \in (\beta, \mu^*)$, the perturbed partition involving pooling for types close enough to zero but strictly positive improves (notice that I've show feasibility (IC)). Now, to see that this is an improvement compared with the full information partition, observe that losses come from a range with measure $\left[F(\epsilon_1) - \frac{1}{2}\right]$, whereas gains come from a measure $[1 - F(\epsilon_1)]$. Hence we have to choose ϵ_1 so as to maximize

$$[1 - F(\epsilon_1)]\, c(\mu^{*2} - \mu^2)\epsilon_1 - \left[F(\epsilon_1) - \frac{1}{2}\right] L(\epsilon_1)$$

First Order Condition for maximization is

$$-f(\epsilon_1)\left[c(\mu^{*2} - \mu^2)\epsilon_1 + L(\epsilon_1)\right] + [1 - F(\epsilon_1)]\left[c(\mu^{*2} - \mu^2) + \epsilon_1 c\frac{\partial \mu}{\partial \epsilon_1}\right] - \left[F(\epsilon_1) - \frac{1}{2}\right] L'(\epsilon_1) = 0$$

Whenever $\epsilon_1 \to 0$ this is clearly positive since it becomes

$$\frac{1}{2}c(\mu^{*2} - \mu^2) > 0$$

□

PROOF OF PROPOSITION 3.7. As before the analysis is made for the case in which $\epsilon > 0$ being the other case equivalent. The proof is by contradiction. Assume first that the result does not hold, then there exist a triple of cutoffs $\{\epsilon_{k-1}, \epsilon_k, \epsilon_{k+1}\}$, with $\epsilon_{k+1} > \epsilon_k > \epsilon_{k-1} > 0$ and such that $m(a) = \hat{m}$ for all $a \in [a^*(\epsilon_k), a^*(\epsilon_{k+1}))$and $m(a) = a$ if $a \in [a^*(\epsilon_{k-1}), a^*(\epsilon_k))$, where $a^*(\epsilon)$ are the original equilibrium actions. Now, take $\hat{\epsilon}$ such that

$$\hat{\epsilon} = \epsilon_{k-1} + (\epsilon_{k+1} - \epsilon_k)$$

I offer a scheme that improves for the principal by setting

$$m(a) = \hat{m} \text{ for all } a \in [a(\epsilon_{k-1}), \tilde{a}(\hat{\epsilon}))$$

and

$$m(a) = a \text{ for all } a \in [\tilde{a}(\hat{\epsilon}), \tilde{a}(\epsilon_{k+1}))$$

where $\tilde{a}(\epsilon)$ are the new equilibrium actions. Notice that, by construction and linearity of the best responses.

$$\tilde{a}(\epsilon_{k+1}) = a^*(\epsilon_{k+1})$$

Therefore we have

$$a^{FB}(\epsilon) \geq \tilde{a}(\epsilon) \geq a^*(\epsilon_{k+1})$$

for all $\epsilon > 0$, with the second inequality being strict for all $\epsilon \in (\epsilon_{k-1}, \epsilon_{k+1})$. Finally notice that the amount of information is the same in both cases, so that the mechanism is clearly an improvement. □

PROOF OF PROPOSITION 3.8. For the first part, suppose not. Then there exists $\lambda_1 \geq \lambda_2$ such that $\epsilon^*(\lambda_1) < \epsilon^*(\lambda_2)$. There must be a k_1 such that

$$k_1 = \min_k \{a_k(\lambda_1) < a_k(\lambda_2)\}$$

I show that the partition generated by substituting $\hat{a}_k(\lambda_1) = a_k(\lambda_2)$ and leaving the rest unchanged improves upon the original partition. Notice that, since it was optimal before, the marginal loss from worse information equated the marginal gain from better actions at that point. Clearly, the marginal loss from worse information is lower now, since the need for coordination is higher. Hence, all we have to show is that the marginal gain from better actions is not smaller whenever the coordination need is higher.

For the second part, the argument is similar, noticing that higher variance means lower importance of small shocks, as compared to big shocks and since the advantage of pooling up to type ϵ^* is proportional to $1 - F(\epsilon^*)$, higher variance leads to ϵ^* being higher. $\square$

Bibliography

Acemoglu, D., K. Bimpikis, and A. Ozdaglar (2010, September). Dynamics of information exchange in endogenous social networks. NBER Working Papers 16410, National Bureau of Economic Research, Inc.

Allen, F., S. Morris, and H. S. Shin (2006). Beauty contests and iterated expectations in asset markets. *Review of Financial Studies 19*(3), 719–752.

Alonso, R., W. Dessein, and N. Matouschek (2008). When does coordination require centralization? *American Economic Review 98*(1), 145–79.

Ambrus, A. and G. Egorov (2009). Delegation and nonmonetary incentives. *mimeo*.

Anderlini, L., D. Gerardi, and R. Lagunoff (2012). Communication and learning. *Review of Economic Studies*.

Andrew, C. (2009). *The Defense of the Realm: The Authorized History of MI5*. London: Allen Lane.

Aral, S., E. Brynjolfsson, and M. V. Alstyne (2007, June). Information, technology and information worker productivity: Task level evidence. NBER Working Papers 13172, National Bureau of Economic Research, Inc.

Arrow, K. (1974). *The Limits of Organization*. The Fels Lectures on Public Policy Analysis. New York: W.W.Norton and Company.

Athey, S. and J. Roberts (2001, May). Organizational design: Decision rights and incentive contracts. *American Economic Review 91*(2), 200–205.

Bar-Isaac, H. (2003, 04). Reputation and survival: Learning in a dynamic signalling model. *Review of Economic Studies 70*(2), 231–251.

Benabou, R. and G. Laroque (1992, August). Using privileged information to manipulate markets: Insiders, gurus, and credibility. *The Quarterly Journal of Economics 107*(3), 921–58.

Bergemann, D. and J. Valimaki (2002, May). Information acquisition and efficient mechanism design. *Econometrica 70*(3), 1007–1033.

Blackwell, D. W. and D. B. Winters (1997, Summer). Banking relationships and the effect of monitoring on loan pricing. *Journal of Financial Research 20*(2), 275–89.

Bloom, N., R. Sadun, and J. V. Reenen (2012, February). Americans do it better: Us multinationals and the productivity miracle. *American Economic Review 102*(1), 167–201.

Bolton, P. and M. Dewatripont (1994, November). The firm as a communication network. *The Quarterly Journal of Economics 109*(4), 809–39.

Bresnahan, T. F., E. Brynjolfsson, and L. M. Hitt (2002, February). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics 117*(1), 339–376.

Caballero, R. J., T. Hoshi, and A. K. Kashyap (2008, December). Zombie lending and depressed restructuring in japan. *American Economic Review 98*(5), 1943–77.

Calvó-Armengol, A., J. de Martí, and A. Prat (2011). Endogenous communication in complex organizations. *mimeo*.

Chandler, A. D. (1969, June). *Strategy and Structure: Chapters in the History of the American Industrial Enterprise*, Volume 1 of *MIT Press Books*. The MIT Press.

Crawford, V. P. and J. Sobel (1982, November). Strategic information transmission. *Econometrica 50*(6), 1431–51.

Dessein, W., L. Garicano, and R. Gertner (2010, November). Organizing for synergies. *American Economic Journal: Microeconomics 2*(4), 77–114.

Dessein, W. and T. Santos (2006, October). Adaptive organizations. *Journal of Political Economy 114*(5), 956–985.

Dewatripont, M. and E. Maskin (1995, October). Credit and efficiency in centralized and decentralized economies. *Review of Economic Studies 62*(4), 541–55.

Diamond, D. W. (1991, August). Monitoring and reputation: The choice between bank loans and directly placed debt. *Journal of Political Economy 99*(4), 689–721.

Eccles, R. (1985). *The transfer pricing problem: a theory for practice*. Lexington books. Lexington Books.

Ely, J. C. and J. Valimaki (2003, August). Bad reputation. *The Quarterly Journal of Economics 118*(3), 785–814.

Fama, E. F. (1985). What's different about banks? *Journal of Monetary Economics 15*(1), 29–39.

Galeotti, A., C. Ghiglino, and F. Squintani (2009, May). Strategic information transmission in networks. Economics Discussion Papers 668, University of Essex, Department of Economics.

Galeotti, A. and S. Goyal (2010, September). The law of the few. *American Economic Review 100*(4), 1468–92.

Garcia-Gonzalez, D. (2012). Reputation with endogenous monitoring. Technical report.

Garicano, L. (2000, October). Hierarchies and the organization of knowledge in production. *Journal of Political Economy 108*(5), 874–904.

Garicano, L. and R. A. Posner (2005, Fall). Intelligence failures: An organizational economics perspective. *Journal of Economic Perspectives 19*(4), 151–170.

Geanakoplos, J. and P. R. Milgrom (1988, October). A theory of hierarchies based on limited managerial attention. Cowles Foundation Discussion Papers 775R, Cowles Foundation for Research in Economics, Yale University.

Hagenbach, J. and F. Koessler (2010, 07). Strategic communication networks. *Review of Economic Studies 77*(3), 1072–1099.

Hansen, S. E. (2010, July). The benefits of limited feedback in organizations. Economics Working Papers 1232, Department of Economics and Business, Universitat Pompeu Fabra.

Hellwig, C. and L. Veldkamp (2009, 01). Knowing what others know: Coordination motives in information acquisition. *Review of Economic Studies 76*(1), 223–251.

Hermalin, B. E. (1998, December). Toward an economic theory of leadership: Leading by example. *American Economic Review 88*(5), 1188–1206.

Hertzberg, A., J. M. Liberti, and D. Paravisini (2010, 06). Information and incentives inside the firm: Evidence from loan officer rotation. *Journal of Finance 65*(3), 795–828.

Holmstrom, B. (1980, June). On the theory of delegation. Discussion Papers 438, Northwestern University, Center for Mathematical Studies in Economics and Management Science.

Holmstrom, B. (1999). Managerial incentive problems: A dynamic perspective. *The Review of Economic Studies 66*(1), 169–182.

Hong, H., J. D. Kubik, and J. C. Stein (2005, December). Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers. *Journal of Finance 60*(6), 2801–2824.

Ivanov, M. (2010, March). Informational control and organizational design. *Journal of Economic Theory 145*(2), 721–751.

Kandori, M. and I. Obara (2005). Endogenous monitoring. *Unpublished Manuscript*.

Kano, M., H. Uchida, G. F. Udell, and W. Watanabe (2011, April). Information verifiability, bank organization, bank competition and bank-borrower relationships. *Journal of Banking & Finance 35*(4), 935–954.

Kreps, D. M. and R. Wilson (1982, August). Reputation and imperfect information. *Journal of Economic Theory 27*(2), 253–279.

Lee, D. and E. V. den Steen (2010, February). Managing know-how. *Management Science 56*(2), 270–285.

Liu, Q. (2011). Information acquisition and reputation dynamics. *Review of Economic Studies forthcoming*.

Mailath, G. and L. Samuelson (2006). *Repeated games and reputations: long-run relationships*. Oxford University Press.

Marschak, J. and R. Radner (1972). *Economic theory of teams [by] Jacob Marschak and Roy Radner*. Yale University Press, New Haven,.

Milgrom, P. and J. Roberts (1982, August). Predation, reputation, and entry deterrence. *Journal of Economic Theory 27*(2), 280–312.

Morris, S. (2001, April). Political correctness. *Journal of Political Economy 109*(2), 231–265.

Morris, S. and H. S. Shin (2002, December). Social value of public information. *American Economic Review 92*(5), 1521–1534.

Myatt, D. P. and C. Wallace (2011). Endogenous information acquisition in coordination games. *The Review of Economic Studies*.

Ozsoylev, H. N. and J. Walden (2011). Asset pricing in large information networks. *Journal of Economic Theory 146*(6), 2252 – 2280.

Podolny, J., J. Roberts, and A. Kemper (2000, July). Novo nordisk (a) global coordination. Case study, Stanford GSB.

Radner, R. (1993). The organization of decentralized information processing. *Econometrica 61*(5), pp. 1109–1146.

Raith, M. (1996, October). A general model of information sharing in oligopoly. *Journal of Economic Theory 71*(1), 260–288.

Rantakari, H. (2008, October). Governing adaptation. *Review of Economic Studies 75*(4), 1257–1285.

Roberts, J. (2004). *The modern firm: organizational design for performance and growth*. Clarendon Lectures in Management Studies. Oxford University Press.

Shiller, R. J. and J. Pound (1986, March). Survey evidence on diffusion of investment among institutional investors. NBER Working Papers 1851, National Bureau of Economic Research, Inc.

Sobel, J. (1985, October). A theory of credibility. *Review of Economic Studies 52*(4), 557–73.

Szalay, D. (2005, October). The economics of clear advice and extreme options. *Review of Economic Studies 72*(4), 1173–1198.

Xia, C. (2010). Communication and confidence in financial networks. *mimeo*.